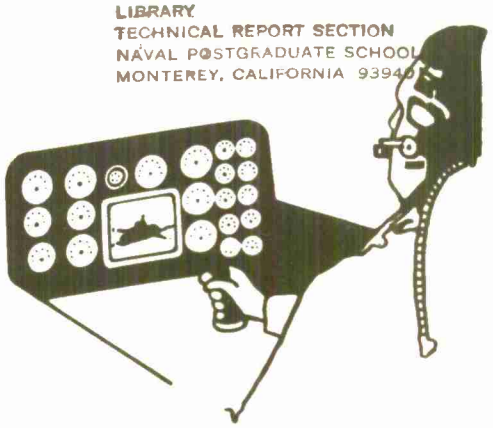


VOLUME I

# A COLLECTION OF UNCLASSIFIED TECHNICAL PAPERS ON TARGET ACQUISITION

OA 6201

ONR Contract No. N000014-72-C-0389 Work Unit No. NR 196-121

PREPARED BY:
MARTIN MARIETTA AEROSPACE
ORLANDO DIVISION

Approved for Public Release;
Distribution Unlimited.



## OFFICE OF NAVAL RESEARCH

# TARGET ACQUISITION SYMPOSIUM

## NAVAL TRAINING CENTER, ORLANDO, FLORIDA / 14,15,16 NOVEMBER 1972

Unclassified

Security Classification

## DOCUMENT CONTROL DATA - R & D

*(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)*

| 1. ORIGINATING ACTIVITY (Corporate author) | 2a. REPORT SECURITY CLASSIFICATION |
|---|---|
| Martin Marietta Aerospace<br>Orlando, Florida | Unclassified |
| | 2b. GROUP N/A |

3. REPORT TITLE

A Collection of Unclassified Technical Papers on Target Acquisition. Volume I

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Technical Report

5. AUTHOR(S) (First name, middle initial, last name)

Daniel B. Jones, Editor
Martin Marietta Aerospace

| 6. REPORT DATE | 7a. TOTAL NO. OF PAGES | 7b. NO. OF REFS |
|---|---|---|
| March 1973 | 228 | See Individual Papers |

| 8a. CONTRACT OR GRANT NO. | 9a. ORIGINATOR'S REPORT NUMBER(S) |
|---|---|
| N000014-72-C-0389 | OA 6201 Volume I |
| b. PROJECT NO. Work Unit Number NR 196-121 | |
| c. | 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) N/A |
| d. | |

10. DISTRIBUTION STATEMENT

Approved for public release; distribution unlimited

| 11. SUPPLEMENTARY NOTES | 12. SPONSORING MILITARY ACTIVITY |
|---|---|
| N/A | Office of Naval Research, Code 455<br>Arlington, Virginia 22217 |

13. ABSTRACT

This report is the Unclassified Volume of the Proceedings of the Target Acquisition Symposium which was held at the Naval Training Center, Orlando, Florida on 14, 15, 16 November 1972, sponsored by the Engineering Psychology Programs, Office of Naval Research, Arlington, Virginia and conducted by the Martin Marietta Aerospace, Orlando, Florida.

The symposium presented current research in the field of target acquisition and included a panel discussion by pilots illustrating operational problems.

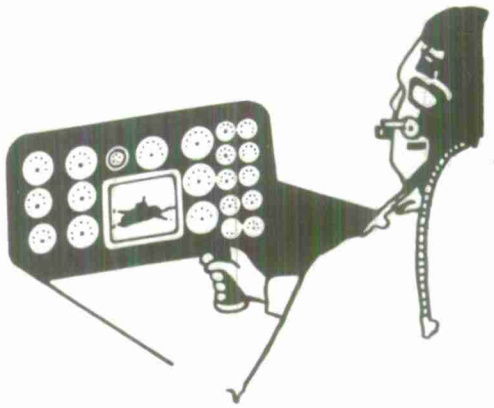
The classified technical papers are presented in Volume II, OA 6201, entitled, ''A Collection of Classified Technical Papers on Target Acquisition (U).''

DD FORM 1 NOV 65 1473

Unclassified

Security Classification

| 14 KEY WORDS | LINK A | | LINK B | | LINK C | |
|---|---|---|---|---|---|---|
| | ROLE | WT | ROLE | WT | ROLE | WT |
| Symposium | | | | | | |
| Target Acquisition | | | | | | |
| Search Radar | | | | | | |
| Laser Target Designator | | | | | | |
| Infrared Signatures | | | | | | |
| FLIR | | | | | | |
| Low Light Level TV | | | | | | |
| Electro Optical Devices | | | | | | |
| Visual Search | | | | | | |
| Math Modelling | | | | | | |
| Optical Countermeasures | | | | | | |
| Target Cueing | | | | | | |
| Imaging Devices | | | | | | |
| Terrain Simulator | | | | | | |



VOLUME I

# A COLLECTION OF UNCLASSIFIED TECHNICAL PAPERS ON TARGET ACQUISITION

OA 6201

ONR Contract No. N000014-72-C-0389 Work Unit No. NR 196-121

PREPARED BY:

MARTIN MARIETTA AEROSPACE
ORLANDO DIVISION



## OFFICE OF NAVAL RESEARCH

## TARGET ACQUISITION SYMPOSIUM

NAVAL TRAINING CENTER, ORLANDO, FLORIDA / 14,15,16 NOVEMBER 1972

## FOREWORD

The Target Acquisition Symposium held at the Naval Training Center, Orlando, Florida, on 14, 15, 16 November 1972 was sponsored by the Engineering Psychology Programs, Office Of Naval Research, Washington, D. C. and was conducted by Martin Marietta Aerospace, Orlando, Florida under ONR contract No. N000014-72-C-0389. The proceedings are contained in two volumes:

Vol I. A Collection of Unclassified Papers on Target Acquisition

Vol II. A Collection of Classified Papers on Target Acquisition (U)

The papers appear in each volume in chronological order as presented. The symposium was held in six sessions moderated as follows:

Session 1 – Dr. James J. Regan
2 – Dr. Antonio Mendez
3 – Dr. Harry L. Snyder
4 – Dr. Daniel B. Jones
5 – Dr. John J. O'Hare
6 – Dr. Martin A. Tolcott

A brief biographical sketch of each Panel Moderator, Panel Participant, and Author is included at the end of each volume.

Special acknowledgement is made for the personal sponsorship of the Symposium by Dr. Martin A. Tolcott, ONR, and to James L. Carpenter and Kathryn W. Leonard of Martin Marietta for the difficult tasks of preparing for and conducting the symposium.

Daniel B. Jones, Editor

## VOLUME I

CONTENTS



The following papers presented at the Target Acquisition Symposium have been published in Volume II, ''A Collection of Classified Technical Papers on Target Acquisition (U)'' of OA 6201.

1. Condor Performance in an Operational Environment: a Field Test and Mathematical Model Validation.
2. Aircrew Target Acquisition Problems, a Panle Discussion.
3. Optical Countermeasures Considerations.
4. Target Cueing System for Far IR Fire Control Systems.
5. Tactical Target Acquisition with Laser Protective Visors.
6. Real-Time Interpretation of Infrared Imagery.
7. A Description of Air-to-Ground Target Acquisition in England and Germany.
8. High Resolution Side-Looking Ground Mapping Radar and Operator Performance.

TARGET ACQUISITION SYMPOSIUM
OPENING REMARKS

by

Fred A. Payne
Vice President and Chief Engineer
Martin Marietta Aerospace
Orlando, Florida

# OFFICE OF NAVAL RESEARCH
# TARGET ACQUISITION SYMPOSIUM
## NAVAL TRAINING CENTER, ORLANDO, FLORIDA / 14,15,16 NOVEMBER 1972

TARGET ACQUISITION SYMPOSIUM
OPENING REMARKS
by
Fred A. Payne
Vice President and Chief Engineer
Martin Marietta Aerospace
Orlando Division

PROBLEM

The problem of accurately detecting and recognizing targets with either the unaided eye, or even with the use of sophisticated sensing devices, seriously limits the design and operation of reconnaissance, fire control, and air launched missile systems. The state-of-the-art in missile design and guidance systems has far outstripped the ability of the human operator to acquire ground targets at maximum launch ranges. The requirements for low level high speed attack profiles also compounds the crew's difficulty in acquiring targets at safe distances. For example, current optical correlator weapons systems are limited in the acquire-before-launch mode by the pilot's ability to detect targets. If the potential of sensor and missile technology is to be exploited, real time target detection and recognition will have to be improved by means of more sophisticated sensing/display systems and the associated man/display interfaces.

The advent of "Smart" weapons has made the target acquisition problem very real. Today we can accurately guide tactical weapons to small, hard targets. Many of these weapons can be released at slant ranges beyond 50,000 feet. Guidance accuracies of less than five feet can be achieved with CEP's. Direct hits are being achieved on tanks or trucks from ranges well beyond any previous capability. Even more accurate, longer range, and more diverse weapons are in development. But our detection systems still do not match our weapons capability.

For years we have relied on the pilot's or observer's unaided visual capability to acquire targets. But the human "Mark I eyeball" has not been improved. Unassisted eye detection is physiologically limited. Under ideal conditions our reliable visual target acquisition performance will be marginal. Experimental evidence shows that, even on a clear day, a static target twenty feet long with high contrast will be found by an observer flying at 400 knots speed at ten thousand feet altitude with a high probability of detection at less than 20,000 feet slant range. Deduct aircraft trim up and weapon release times from that distance and we are well inside the range of accurate enemy anti-aircraft weapons. If the target is really "tactical" - small, mobile, and camouflaged - then we probably will miss it entirely.

Our current operational family of electro-optical and electronic sensors do not significantly improve the situation. While we can detect some targets with TV, FLIR's or Radar at greater than eyeball ranges, these longer range systems are restricted by narrow fields of view and limiting resolution. None of them approach the Mark I eyeball in resolution or in field of view. If we know where to point the system, then we may have a high probability of target acquisition out to perhaps 40 thousand feet. This still means that the real tactical targets are usually missed completely, since only a small portion of the potential target area is viewed.

Our dilemma can be graphically illustrated. Typical air-to-ground weapon performance is shown in Figure 1. Target acquisition, even by our best sensor equipment and weapon release, means the pilot will be flying in the area marked acquisition (Figure 1(a)).

Even small anti-aircraft weapons fire out to nearly 20,000 feet and medium AA guns to around 40,000 feet (Figure 1(b)).

If we add enemy SAM's to this situation the attacking aircraft is subject to fire long before he can possibly acquire the target (Figure 1(c)). Our weapons capability will allow us to attack from relatively safe distances, but we cannot find the targets.

The costs of this problem are very real. We can successfully destroy tactical targets at relatively safe ranges, but our

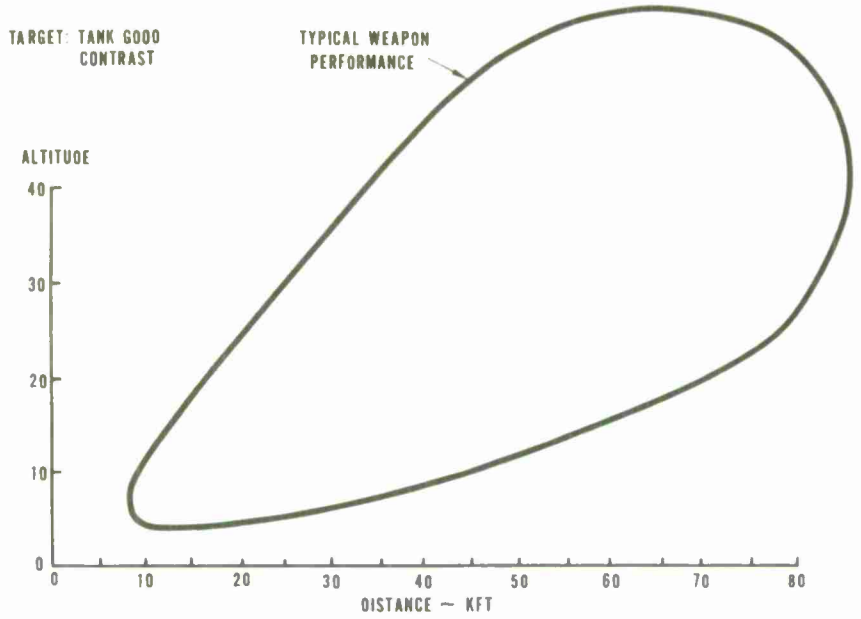



Figure 1. Target Acquisition's Problem

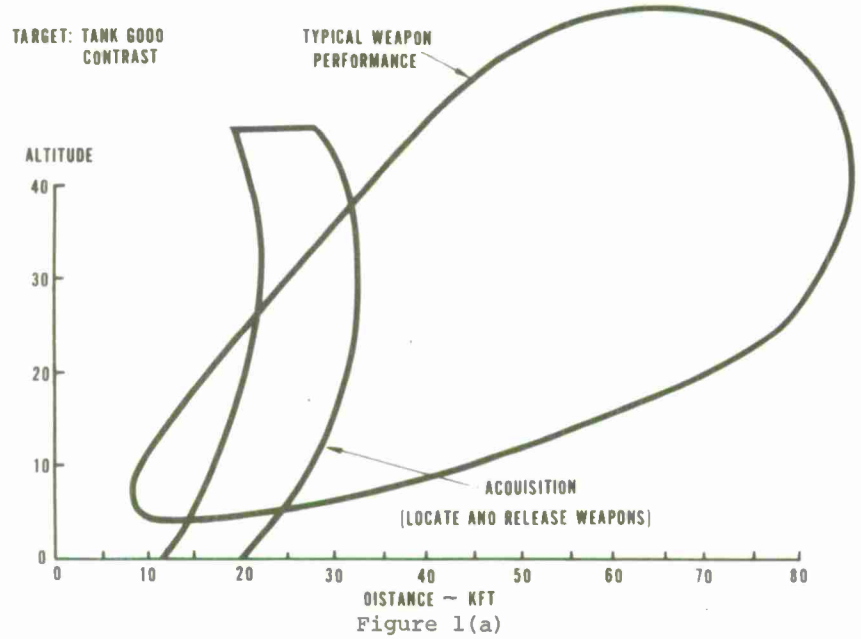



Figure 1(a)

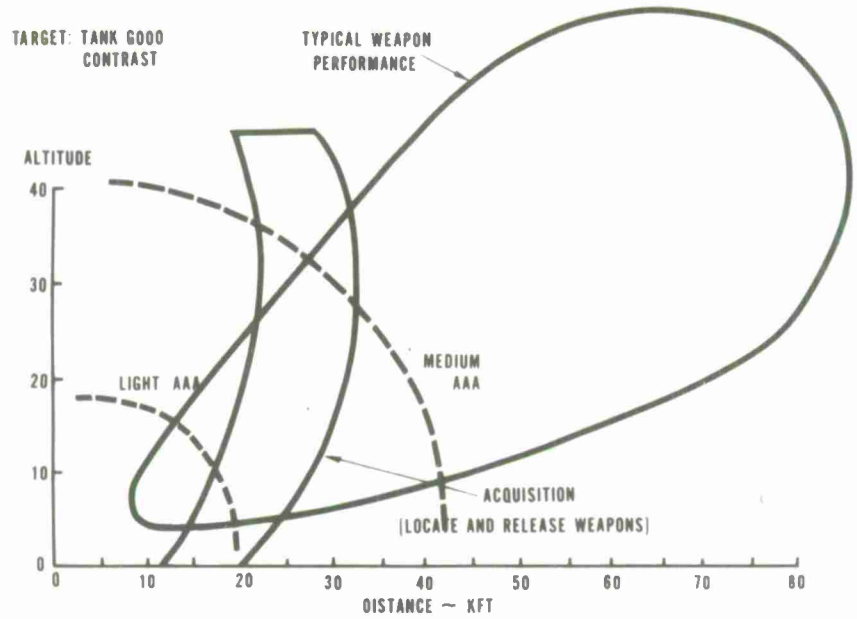



Figure 1(b)

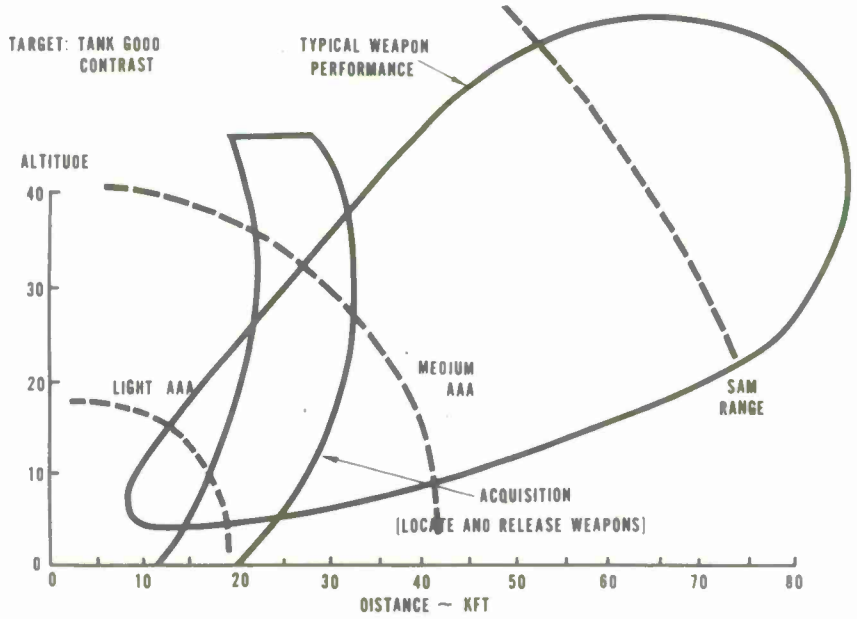



Figure 1(c)

acquisition technology does not match. In Viet Nam we still rely upon visual target acquisition for tactical targets. This causes the attrition rate for the Forward Air Control aircraft, which are used to acquire targets, to be the highest of any aircraft group in the theater.

This gentlemen is our problem. How can we find targets at ranges commensurate with our weapons' capability?

In the next three days we will be considering some ways to solve that problem. There are four elements in any solution: 1) the sensors; 2) the data processor; 3) displays; 4) the human.

Let's all keep a mental box score during the symposium to determine where the effort is and is not being invested.

OBJECTIVE

The objective of this symposium on target acquisition is first, to assess the state-of-the-art in acquisition technology; second, to establish a dialog among the many groups represented here - from the theoretician to the crew member who ultimately must use the system - and to create an understanding of all of the interrelated problems. We hope also that you may pose some new approaches to solving these problems. What we need are practical solutions to a long standing problem.

The symposium is oriented toward the human factors aspect because target acquisition eventually comes down to the crewman trying to do a job. However, all of us concerned with the problem are represented: pilots, engineers, designers, operations analysts, and human factors specialists.

SCOPE

The scope of the three day session reflects our diversity of interest:

- Operational Techniques,
- Countermeasures to Target Acquisition,
- Sensor Systems,
- Operational and Systems Models, and
- Image and Display Systems.

Our symposium is planned to consider each of these areas from the perspective of our several divergent technical interests. We hope to emphasize real problems and practical solutions.

THANKS

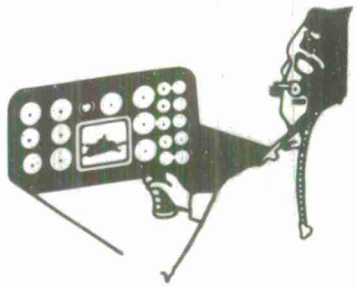
We at Martin Marietta are especially pleased that the Office of Naval Research has shown a continuing interest in these practical solutions. We believe that the applied research in this area, that has been sponsored by ONR, will result in better and more effective target acquisition systems.

We particularly wish to thank:

- Dr. Martin Tollcott, Director of ONR's Engineering Psychology Programs, for his direct interest that made this meeting possible.
- Captain J. F. Gillooly, Commander of the Orlando Naval Training Center, who has graciously provided us with the facilities and support necessary to conduct our meeting.
- Captain A. G. Finley, Commanding Officer of the Naval Training Equipment Center who has lent the services and time of his personnel to assist us.

These gentlemen, and myself, wish to welcome you here today. Find the Target and keep your attention there for these next three days.

I will now turn you over to our Symposium Chairman, Dr. Daniel B. Jones.



# THE EMETF WEAPON SYSTEM ELECTROMAGNETIC ENVIRONMENT SIMULATOR

by

R. L. McCluskey
U.S. Army Electronic Proving Ground
Fort Huachuca, Arizona

OFFICE OF NAVAL RESEARCH
TARGET ACQUISITION SYMPOSIUM
NAVAL TRAINING CENTER, ORLANDO, FLORIDA / 14,15,16 NOVEMBER 1972

# THE EMETF WEAPON SYSTEM ELECTROMAGNETIC ENVIRONMENT SIMULATOR

R. L. McCluskey
U. S. Army Electronic Proving Ground

ABSTRACT

This paper, presented as a slide film briefing, delineates the capabilities and operational modes of a Weapon System Electromagnetic Environment Simulator. This simulator is part of the measurement facilities of the Electromagnetic Environmental Test Facility (EMETF) of the U.S. Army Electronic Proving Ground at Fort Huachuca, Arizona. In this discussion the capabilities and operation of a mathematical interference prediction model, as well as other test facilities of the EMETF, are related to the Weapon System Electromagnetic Environment Simulator. This relation of prediction and simulation software to the empirical testing of component and data links in testing is explained. The development history of the EMETF and plans for a future interactive and integrated test capability conclude the briefing.

The purpose of this unclassified briefing is to acquaint you with the Weapon System Electromagnetic Environment Simulator, the WSEES. In addition, we will discuss the capability of the EMETF to evaluate the electromagnetic compatibility of a total weapon system.

The WSEES is capable of simulating, under controllable and repeatable conditions, the exact radio frequency signals associated with a radar or weapon system. Simultaneously it can simulate signals associated with other weapon systems, ECM devices, and communications-electronics systems that might degrade the performance of the system being evaluated.

This simulator is part of the measurement facilities of the Electromagnetic Environment Test Facility of the U.S. Army Electronic Proving Ground at Fort Huachuca, Arizona, a testing activity of U. S. Army Test and Evaluation Command. The mission of the EMETF is to determine the electromagnetic compatibility – EMC – of Army communications-electronics equipment, systems, and concepts. To place the WSEES in the context of the EMETF's mission, I will give a brief summary of the EMETF's history and the requirements for the WSEES.

''Electromagnetic compatibility'' is defined in AR 11-13. The key words are ''intended operational environments'' and ''unwanted electromagnetic radiation or response''. This includes both intentional and unintentional interference.

Electromagnetic compatibility became an issue because of the Army's increasing use of weapons, radios, and other devices using radio frequencies in a limited portion of the electromagnetic spectrum. A study was conducted in the 1950's to identify any problems that might result from this trend and to devise solutions to such problems. The study was approved by the Secretary of the Army in 1958. It resulted in the establishment, in 1960, of the Electromagnetic Environmental Test Facility – or EMETF.

Initially, EMETF testing was conducted in a field facility. This field facility was to contain a U.S. corps composed of three on-line infantry divisions and a reserve armored division in a defensive posture along the Gila River. Deploying and evaluating the required communications-electronics systems in this facility soon proved to be both costly and time consuming, and was plagued by technical measurement and control problems. These difficulties led to the present concept of operation of the EMETF.

To make the many measurements required by our diverse analyses, the EMETF uses six

interrelated facilities: the Field Facility, where we measure equipments under field conditions; the Instrumented Workshop, where we perform laboratory-type measurements and closed link tests; the Spectrum Signature Facility, where we analyze spectrum characteristics of emitters; the Scoring Facility, where we measure information transfer of communications-electronics (CE) systems; and the WSEES, where we determine the effects of interference on weapon systems and radars (Figure 1). Finally, in the computer-automated Interference Prediction Model, we use data from these five measurement facilities to perform large-scale analyses of CE tactical deployments and operational concepts.

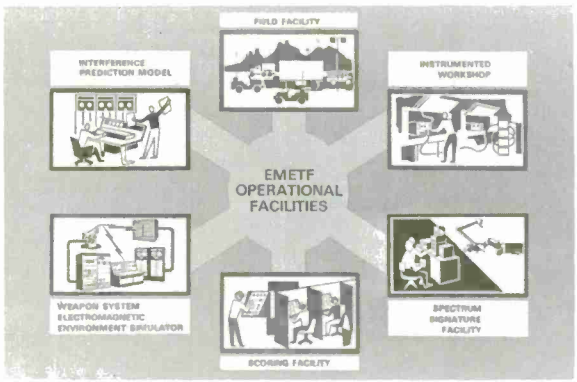



Figure 1. EMETF Operational Facilities

The Interference Prediction Model, or IPM, is a major analysis tool of the EMETF and requires basic types of input data. The test bed provides information on equipment types and locations, equipment linking and netting, and other pertinent data concerning the deployment and operational concept being evaluated. The characteristics of the equipments being evaluated are provided by the spectrum signature data. The propagation model provides propagation path loss data. Scoring data are performance information on a particular equipment that enables us to rate its effectiveness with and without interference or jamming. We determine scoring data by empirical measurements under controlled conditions. The IPM then determines the equipment performance in the tactical environment both with and without interference.

Communications equipment is subject to interference from any signal on the same frequency. The desired signal when combined with a sufficiently powerful interfering signal is likely to become entirely unintelligible. Measuring these relative signal to interference levels in our Instrumented Workshop yields the scoring data needed (Figure 2).

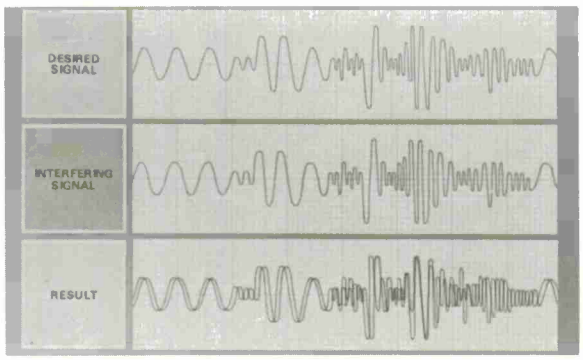



Figure 2. Communications Interference

A radar receiver, however, may not ''notice'' a signal other than its desired signal, even when the two are on the same frequency, if their timing and phasing differ. The upper part of Figure 3 illustrates this type of situation. The lower part shows interfering pulses in time phase with some of the desired signal pulses. When the signals are in time phase, interference may occur. When they are not in time phase, interference may not be noticeable even if the undesired signal is much stronger than the desired signal. Obtaining data for scoring the effectiveness of radars and weapon systems thus require more complex equipment and procedures than for communications equipment.

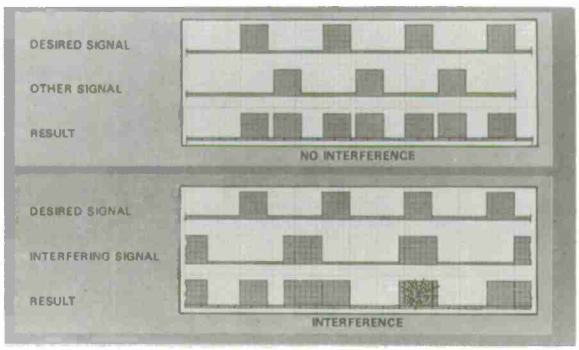



Figure 3. Radar Interference

Prior to development of the Weapon System Electromagnetic Environment Simulator, the WSEES, scoring data on radars and weapon systems had to be obtained by field testing. Not only was this time consuming and laborious, but data were frequently difficult to verify because of the non-repeatable nature of field conditions, and costs were often exorbitant.

We therefore developed a facility similar in concept to the Instrumented Workshop, and capable of controlled and repeatable testing of radars and weapon systems. This simulator, the WSEES, enables us to analyze the EMC of such equipment without elaborate field tests.

The simulator had to have the capability of simulating the RF signals associated with air defense, shell and mortar tracking, and combat surveillance radars. These signals include pulse, CW and both pulse and CW doppler.

Figure 4 shows the overall WSEES including both hardware and software. Since the EMETF uses the Interference Prediction Model or IPM to determine equipment performance in the tactical situation, the primary purpose of the WSEES is to provide data on the effects of interference on radars and weapon systems. Therefore, we can use the general purpose computer at Fort Huachuca to prepare our control tapes. These are then used to control the RF signals generated by the basic simulator.

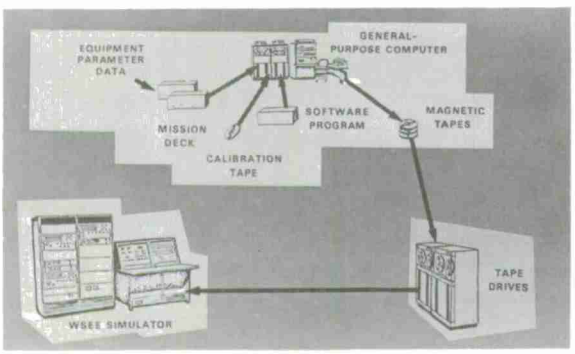



Figure 4. Overall WSEES System

The present WSEES equipment is shown in Figure 5. On the right are the tape input units; in the center, the manual control unit and digital interface. The two racks on the left contain the RF generation equipments.

Figure 5. Present WSEES Equipment

The digitally controlled RF generator section contains three RF generators, each operating in three bands from 2 to 12.4 GHz (Figure 6). Two of these generators may be phaselocked to the equipments under test to provide coherent signals. Each generator can be time-shared to simulate several different emitters. This permits simulating up to 32 pulsed emitters. The doppler shift range of 30 to 100 kHz allows simulation of a range of targets from a walking man to a mach 9 missile.

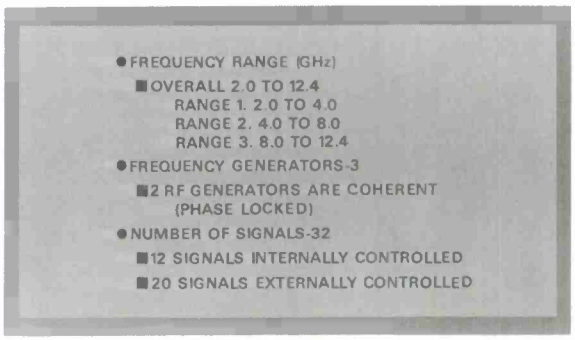



Figure 6. WSEES Characteristics

The WSEES has a highly versatile pulse generation capability. Pulses can have widths from 0.1 to 200 microseconds at repetition rates from 20 to 50,000 pulses per second. The individual pulses can be controlled in rise time, fall time, and pulse width to simulate various signals in the environment.

Additional characteristics are shown in Figure 7. The WSEES is capable of producing chirped pulses and pulse burst patterns as indicated. Because the present WSEES contains only three RF generators, it is limited to producing three continuous wave or CW signals.

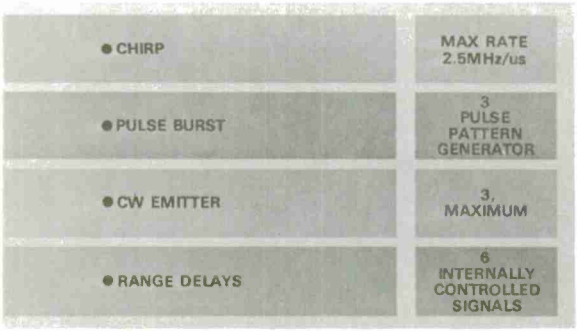



Figure 7. WSEES Special Features

A functional diagram of RF generation control is shown in Figure 8. The coarse frequency control is applied to the cathode of the backward wave oscillator (BWO) and the fine frequency control to the helix. The fine frequency control also controls the chirp rate of a pulse, if required. The switch network between the BWO and the traveling wave tube (TWT) provides the pulse width control. The TWT amplifies the signal and also provides the pulse shape control. Finally, the attenuators set the specified power level.

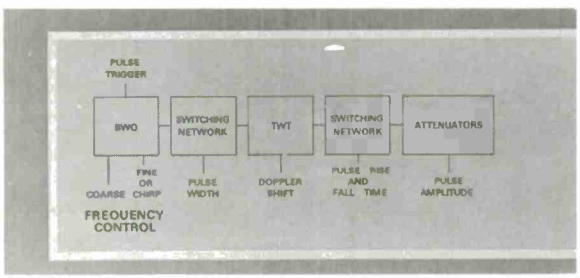



Figure 8. RF Generation Control

Figure 9 is a block diagram of the Phase I WSEES. The RF generation control just discussed is shown in the two blocks labeled ''RF Generators'' and ''Power Controls.'' The control information is provided to the digital interface by the manual input device or the tape units. The digital interface stores information on as many as 32 emitters, and passes these data to the RF generators when needed. The digital control is shown by the thin lines and the RF signal by the heavy line.

Figure 10 depicts a typical situation that could be simulated by the WSEES. This problem also demonstrates the versatility of the system. The test radar system, in the foreground is tracking the target aircraft and controlling the missile, which is in flight to intercept the target. The target has fired a decoy forward to confuse the radar and missile and is also using ECM against both the radar and missile. A background radar is also tracking the target, but has not launched an intercept missile.

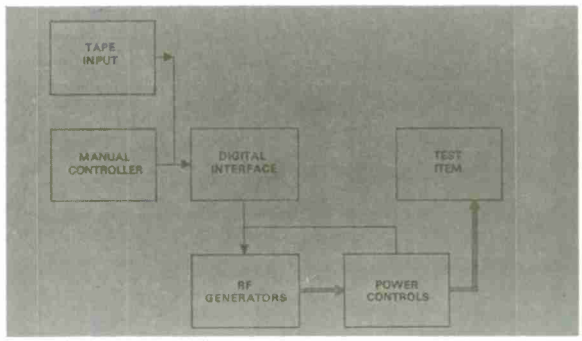



Figure 9. Phase I WSEES

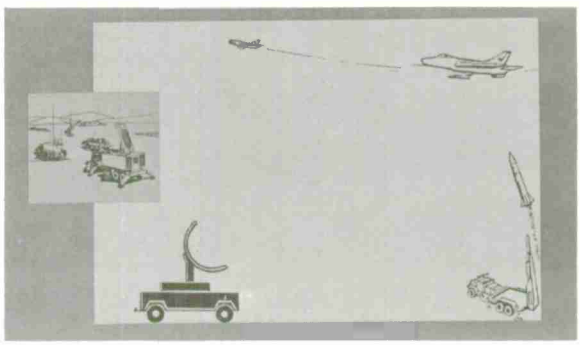


Figure 10. A Typical Problem

The signals to be generated include those reflected to the radar from the target and the control signal from the radar to the missile.

Possible interfering signals from the background radar are those reflected from both target and missile to the test radar and also from target to missile. There may also be a direct signal path from the second radar to the first.

We must also consider the use of an electronic countermeasures system, or systems, against both the radar and the missile.

Combining all these presents a complex electromagnetic environment that can be expected in a typical situation. The WSEES can simulate all of these signals. In addition, it can simulate other signals in the environment that could cause interference.

Future weapon systems may have a multiple reaction capability. They will be able to adapt to interference or jamming by changing frequencies and control modes and other ECCM techniques. In the example shown in Figure 11, if the missile were using semi-active homing, it could lock on and track the decoy, and would therefore miss the target. However, the ground radar by tracking the target, the missile, and the decoy could ascertain that the missile was veering away from the target. The ground unit could then provide guidance commands to bring the missile back on target.

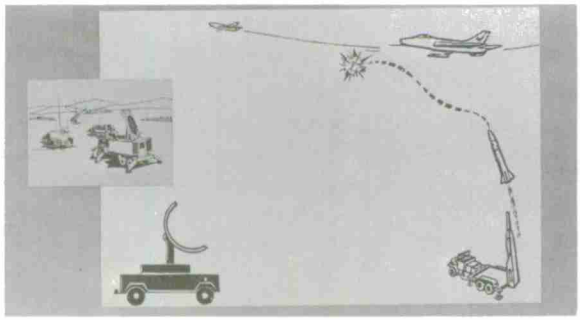


Figure 11. A Typical Problem

To permit rapid testing of these advanced systems, a Phase II development effort is underway to add a real-time control, a steerable platform and other features in Figure 12. I will discuss the application of each of these improvements.

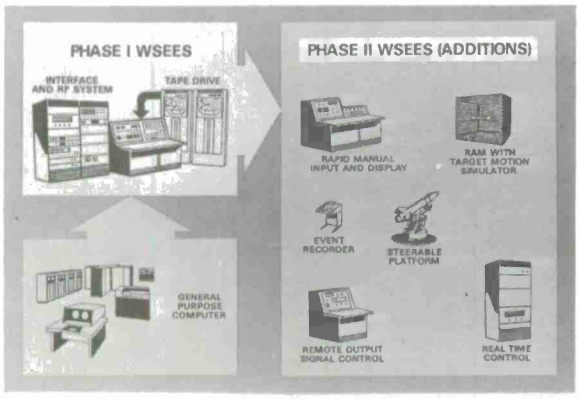



Figure 12. Phase II Development Effort

The real time controller is the heart of the system (Figure 13). It operates in a feed-back loop with the system under test. In its normal mode of operation, the WSEES would generate a signal environment as commanded by the precalculated tapes for the operational situation being simulated. When the operational mode for the system under test changes, the real-time controller would sense this change and alter, as appropriate, the signals being generated. To accomplish the signal changes or modifications required, the real-time controller must not only sense that a change in the operational situation has occurred, but must adjust the signals to reflect the change.

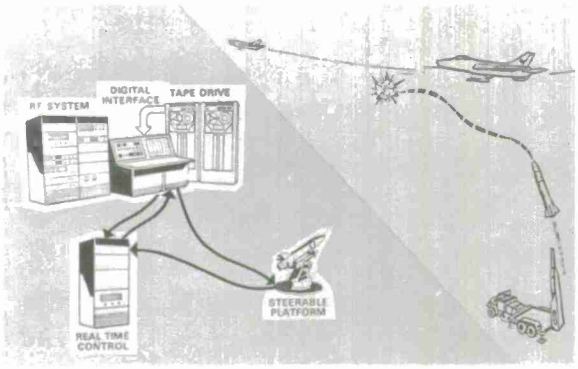



Figure 13. The Real Time Controller

The steerable platform permits simulation and control of the spatial position of the missile. By monitoring the signals to the missile control surfaces in respect to the calculated space position and speed of the missile, the real-time controller can change simulation parameters to reflect changes in missile position.

Four additional items are shown in Figure 14 – a remote output signal control, a RAM enclosure, an event recorder, and a rapid manual input and display.

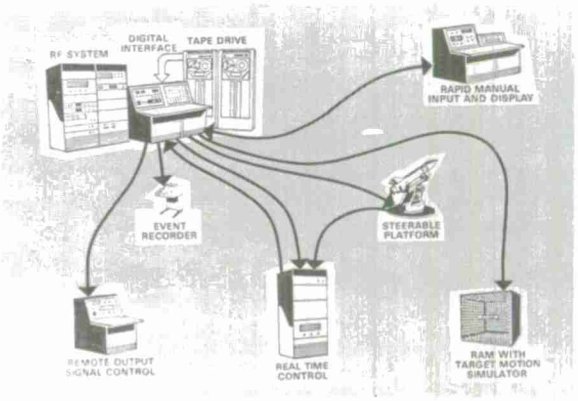



Figure 14. A Remote Output Signal Control, the RAM, and the Event Recorder

The remote output signal control will provide four independently controlled outputs for each of the three RF channels. This will permit direct cabling to the four inputs of a monopulse system, and also increase the number of signal output ports from three to twelve.

The radiation absorbent material enclosure, the RAM, will house radiating antennas which are required for tests when we can't direct wire into the receiver.

The event recorder provides a record of all test conditions and any changes made.

The rapid manual input and display will allow the test situation to be entered manually and changes made by the test engineer without computer generated test tapes.

Figure 15 is an artists's conception of the control panel for the manual input. The initial parameter data for each emitter can be entered by tape, punched card, or manually. The operator can check all data by manually selecting each emitter and observing the parameter data for that emitter. The emitter number is displayed in the center window and the parameter data in the windows on the left. By depressing the appropriate plus or minus button associated with each parameter display, the operator can manually change that parameter as desired.

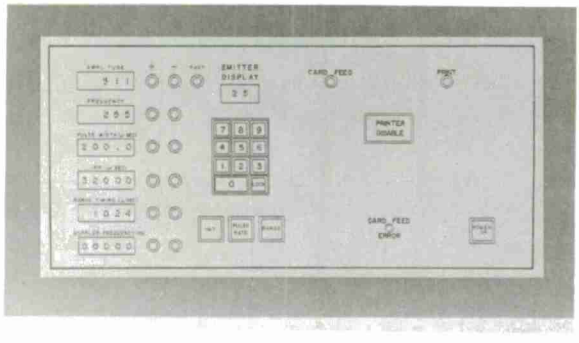



Figure 15. Rapid Manual Input Controller Panel

The Weapon System Electromagnetic Environment Simulator was designed primarily to evaluate the EMC of that portion of a weapon system utilizing radar type signals. However, a total weapon system has many other elements.

To evaluate the voice and data communications of this overall system, the EMETF is currently installing a new Automatic Data Collection System which can automatically test every element of the voice and data transfer system associated with an air defense system.

The heart of the automatic data collection system is the digitally controlled automatic measuring equipment. With either manual or taped control, these equipments can automatically measure and record all of the parameter and test data associated with a communications test link, to include the effects of interference or jamming.

Many future systems may also include so called electro-optical elements and laser communications as elements of an air defense system.

To evaluate these systems the EMETF is in the final design phase of an electro-optical test facility. This facility will analyze the performance of these and other sensors in the EO spectral region.

In summary, the EMETF found that field measurements of military communications-electronics systems were prohibitively slow, variable, and expensive (Figure 16). We developed the WSEES to simulate radar and weapon system emissions. Measurements made from WSEES simulations are fast, repeatable, controllable, and inexpensive. We can measure the effects of these emissions on communications-electronics equipments and gain reliable scoring data for EMC analyses.

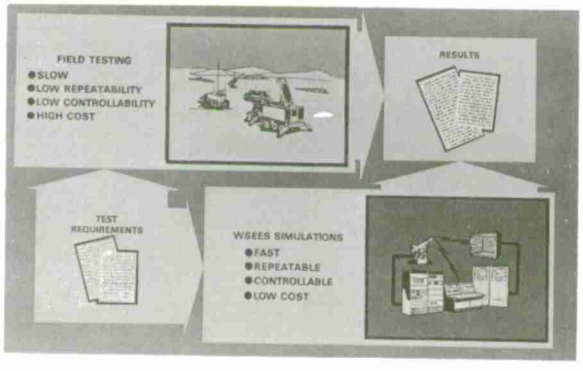
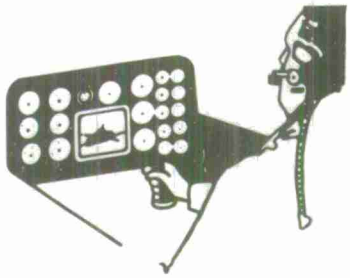



Figure 16. Field Testing Versus Simulation



# AIR-TO-GROUND AND GROUND-TO-AIR DETECTION EXPERIMENTS

by

Dr. Marion R. Bryson
Scientific Advisor, U.S. Army Combat
Developments Experimentation Command
Fort Ord, California 93941

OFFICE OF NAVAL RESEARCH
TARGET ACQUISITION SYMPOSIUM
NAVAL TRAINING CENTER, ORLANDO, FLORIDA / 14,15,16 NOVEMBER 1972

# AIR-TO-GROUND and GROUND-TO-AIR DETECTION EXPERIMENTS

Dr. Marion R. Bryson
Scientific Advisor, US Army Combat
Developments Experimentation Command
Ft Ord, California 93941

## ABSTRACT

Recently, a series of experiments was conducted at Hunter Liggett Military Reservation, California and at Fort Lewis, Washington. The purpose of these experiments was to gain information on the length of time required for a hovering helicopter to detect a ground target and also to determine the time required for the ground element to detect the helicopter. This paper discusses the detection times observed in these experiments and compares them with some theoretical detection curves.

## INTRODUCTION

The US Army Combat Developments Experimentation Command at Ft Ord, California has the mission of performing experiments for the army in which equipment is used by the troops in the field. A series of experiments to assess the capabilities of the attack helicopter is currently being conducted. In 204 experimental trials in this series, time from initiation of line of sight to detection was measured. These detections included both detection of a ground element from a helicopter and detection of the helicopter by the ground element. All of those trials in which time to detect was accurately measured are included in the analyses given in this paper.

Conditions varied considerably from trial to trial. Those factors which were controlled and for which there were an approximately equal number of trials for each level are:

a. Background. The background of the ground element was either cluttered (many trees and bushes although the vehicles were not hidden from view) or open.

b. Movement. The ground element was either moving or stationary.

c. Range. The range from the hovering helicopter to the ground element was either 1, 2 or 3 kilometers.

d. Information. The accuracy of the information the air crew possessed concerning the location of the ground element was either six digit coordinates of its location (exact) or a one square kilometer grid square within which the target was located (1 $km^2$).

e. Crews. There were three different air crews used in the trials.

There were three other controlled factors for which all levels were not equally represented. They are:

(1) On board sighting system/aircraft type. In about two-thirds of the trials, the aircraft was a UH-1B with an XM-26 fire control system aboard. This is the TOW sighting system. The remaining one third were equally divided between two AH-1G helicopters, one with the XM-76 stabilized binoculars and the other with the XM-127 multi-weapons

fire control system.

(2) Aircraft background. The aircraft, while hovering, was silhouetted against the sky less frequently then it had a terrain background.

(3) Geographical location. About two thirds of the trials were at Hunter Liggett Military Reservation in Central California and the remainder were at Fort Lewis, Washington.

## DATA COLLECTION METHODS

A candidate tactic for employment of the attack helicopter is called the pop-up maneuver. In this maneuver, the helicopter crew is given (with variable accuracy) the location of an enemy ground element, generally an armored target, by a forward observer. At the time the air crew receives this information they are 3 to 10 kilometers behind the lines. The helicopter will fly, nap-of-the-earth, toward the enemy until they are within the range of the anti-armor defensive weapon carried by the helicopter. Because of its low altitude and the crews judicious use of cover, it is presumed that the helicopter has remained hidden from the ground targets during this approach. Once in range, the helicopter will come to a hover, behind a terrain, vegetation, or cultural feature which masks it from the enemy. The helicopter will then rise vertically to gain line of sight with the target. It will again come to a hover just above mask while the crew searches for the enemy armored target. Upon detection, the aircraft gunner will fire an anti-armor weapon and, if it is a launcher guided missile, will track the missile to the target. The helicopter will then regain mask as soon as possible.

Although in the trials several time measurements were made, we are interested here only in detection times and total aircraft exposure times. To get an accurate measure of the detection time, the trials were carefully controlled. The ground element, which consisted of five armored vehicles, tanks and armored personnel carriers, was prepositioned on the ground. One helicopter was brought in at nap-of-the-earth to a predesignated launch position, remaining masked from the ground element. On top of the masking position, always a hill, was located a controller with whom the aircraft and another controller with the ground elements had radio communication. The levels of the eight factors discussed above are recorded. The stage is now set for a trial to begin.

The hilltop controller radios the aircraft to unmask. At the time line of sight first exists between the helicopter and the ground elements, as determined by the hilltop controller, he starts two stop watches. The ground controller with the armor force also starts a watch at the time line of sight first exists. When a crew member of the helicopter detects a ground target, he informs the hilltop controller who then stops one of his watches. Similarly, when a crew member of the ground element detects the helicopter, he tells the ground controller. When the helicopter regains mask, the hilltop controller stops his other watch, thereby measuring total exposure time. Safeguards are employed to identify false detects. For the purpose of this paper there are 11 items of information recorded per trial:

a. Range.

b. Aircraft crew.

c. Target acquisition device.

d. Ground target background.

e. Ground target movement.

f. Ground target location information.

g. Aircraft background.

h. Geographic location.

i. Time for aircraft crew to detect.

j. Time for ground to detect the aircraft (if a detection occurred).

k. Total exposure time.

## FACTORS INFLUENCING DETECTION TIMES

When the data from the 204 trials were analyzed, those factors found to have statistically significant effect upon air-to-ground detection time were crew, ground target background and geographic location. Although the significance of these factors is not surprising, one would have expected that range and target movement would have had an effect as well. Such is not the case. The range-information interaction was significant. Six digit coordinate information helped locate the target at the long ranges but was of no value at short ranges.

With regard to ground-to-air detection times (and detection probabilities) range and geographic location were statistically significant factors. There are two surprising results here:

a. That whether the aircraft had a sky or a terrain background made little difference as regards its detectability and

b. That whether the trial was at Hunter Liggett or at Fort Lewis did have a large effect.

A reasonable explanation for both of these results is based on the principle that shape contrast rather than color or reflective contrast may be a more important factor in detection of targets of this nature. The terrain mask above which the aircraft is hovering during the trial is more varied at Ft Lewis, primarily because of heavier vegetation.

## ANALYSIS OF AIR-TO-GROUND DETECTIONS

A common theoretical approach used in search theory assumes that the probability the target is detected during any given unit of time, under the conditions that it has not been detected up to that time, is independent of search time. This leads to the well known equation

$$P_d(t) = 1-(1-P)^t \qquad (1)$$

where

$P_d(t)$ = Probability a detection has occurred in t units of time.

t = Number of units of time search has progressed.

P = Probability a detection occurs in any single unit of time.

The value of the parameter of P can be estimated from empirical data if one assumes that the empirical observations all come from the same population. In the analysis which follows, all 204 observations were combined as if they had a common parametric value. This is a reasonable approach if the assumption is made that the levels of the eight variables discussed in section 1 are randomly chosen and representative of what would occur in actual combat.

A logarithmic transformation of equation (1) yields

$$t\,[\log(1-P)] = \log[1-P_d(t)] \qquad (2)$$

which is linear in t. Fitting the transformed empirical data to this equation, using least squares, yields a value of .062 for P.

Table 1 shows the empirical cumulative probabilities of detection along with the theoretical probabilities

TABLE 1
Probability of Detection: A → G

| Time of Search (t) | Observed Percent of Detections in time (t) | Expected Percent of Detections in time (t) |
|---|---|---|
| 3 | 16 | 17 |
| 6 | 32 | 32 |
| 9 | 48 | 44 |
| 12 | 56 | 53 |
| 15 | 63 | 61 |
| 18 | 66 | 68 |
| 21 | 75 | 74 |
| 24 | 80 | 78 |
| 27 | 82 | 82 |
| 30 | 85 | 85 |
| 33 | 87 | 88 |
| 36 | 90 | 90 |
| 39 | 92 | 92 |
| 42 | 93 | 93 |
| 45 | 94 | 94 |
| 48 | 95 | 95 |

computed using a single second detection probability of .062. Data were grouped into three second increments to eliminate some of the "noise" in the data. The fit is extremely good.

## ANALYSIS OF GROUND-TO-AIR DETECTIONS

The ground-to-air detection data present an interesting problem in data truncation. We possess data on time to detect for those trials in which the ground elements detected the helicopter. All we know about those trials where no detection occurred is that the time to detect for that trial would have been longer than the total aircraft exposure time. Neither the use of only those trials in which detection occurred nor the use of all trials is satisfactory for obtaining a cumulative probability curve.

An unbiased estimate of the probability, $P_d(t)$, the helicopter will be detected in time t can be made. To obtain this estimate, one considers only those trials for which the total aircraft exposure time equalled or exceeded t. If the exposure time was less than t, the trial is ignored, whether there was a ground-to-air detection or not. A completely separate estimate of $P_d(t)$ is made for each time of interest. The estimator is

$$P_d(t) = \frac{D(t)}{E(t)} \qquad (3)$$

where

E(t) = Number of trials for which the aircraft was exposed for time t or greater.

D(t) = Number of the E(t) trials in which the aircraft was detected by the ground in time t or less.

Table 2 shows the values of the factors in equation (3) for selected times. This function need not be monotone increasing with time due to random variation. Under some general conditions, the expected value of $P_d(t)$ is monotone.

TABLE 2
Estimated Probability of Ground-to-Air Detections

| Time(t) | # Aircraft exposed for t seconds: E(t) | # G→A detections: D(t) | $P_d(t)$ |
|---|---|---|---|
| 3 | 204 | 11 | .05 |
| 6 | 204 | 25 | .12 |
| 9 | 202 | 35 | .17 |
| 12 | 189 | 48 | .25 |
| 15 | 167 | 47 | .28 |
| 18 | 149 | 52 | .35 |
| 21 | 128 | 52 | .41 |
| 24 | 104 | 44 | .42 |
| 27 | 89 | 41 | .46 |
| 30 | 77 | 35 | .46 |
| 33 | 65 | 34 | .52 |
| 36 | 54 | 26 | .48 |
| 39 | 44 | 21 | .48 |
| 42 | 35 | 17 | .49 |
| 45 | 29 | 14 | .48 |

It appears from Table 2 that unless the aircraft is detected in about the first 30 seconds of exposure it will not be detected at all. Closer examination of the data reveal two interesting facts.

a. There were several (about 15%) detections in excess of 30 seconds. This would appear to be in direct contradiction to the data shown in Table 2.

b. There is a striking correlation between aircraft exposure time and expected time to detect. This is not, as one might at first think, caused by the fact that a long detection time can only come in a trial where there is a long exposure time. In those trials where the exposure time was 30 seconds or less the aircraft was detected about 53% of the time (detection coming in the first 30 seconds obviously). In those trials where the exposure was greater than 30 seconds the aircraft was detected within 30 seconds in only 37%. This difference becomes even more impressive when one considers that the searcher looked for at least a full 30 seconds in those trials for which exposure exceeded 30 seconds, and had only aircraft exposure time in which to search in the shorter trials. This correlation explains the "flattening" of the ground-to-air detection curve for t greater than 30 seconds.

Fitting the same theoretical curve described in section 4 above to the data shown in the last column of Table 2 yields a poor fit, as expected. At mid-times, around 21 seconds, the curve is about 10 units below the empirical and at 45 seconds it is about 10 units above. Using data from only the first 27 seconds for fitting the curve, an excellent fit is obtained. The empirical and theoretical curves differ by at most two percentage points. The single second probability of detection thus calculated is .023. Table 3 shows the ground-to-air theoretical cumulative probabilities of detection.

## SUMMARY

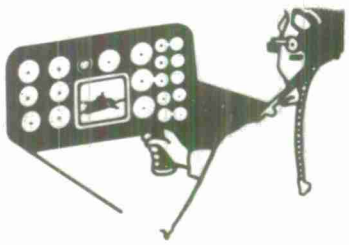
It is felt that the data collected in these experiments are valid in estimating detection probabilities for

TABLE 3

Probability of Detection: G-A

| Time of Search | Observed % of detections in time t | Expected % detections using all data for fitting curve | Expected % detections using only 27 seconds of data for fitting curve |
|---|---|---|---|
| 3 | 5 | 5 | 7 |
| 6 | 12 | 10 | 13 |
| 9 | 17 | 15 | 19 |
| 12 | 25 | 19 | 24 |
| 15 | 28 | 24 | 30 |
| 18 | 35 | 28 | 34 |
| 21 | 41 | 31 | 39 |
| 24 | 42 | 35 | 43 |
| 27 | 46 | 38 | 47 |
| 30 | 46 | 42 | |
| 33 | 52 | 45 | |
| 36 | 48 | 48 | |
| 39 | 48 | 50 | |
| 42 | 49 | 53 | |
| 45 | 48 | 55 | |

random exposures between 1 and 3 kilometers. Both air-to-ground and ground-to-air detection times fit the theoretical curves very well. Using the single second air-to-ground and ground-to-air detection probabilities, .062 and .023, probabilities of first detection were calculated. Theoretically, the air should detect first 72% of the time, the ground first 26% and there should be 2% ties (detection occurring the same second). In the experiment, the air detected first 68% of the time, the ground 28% and 4% were ties. This adds further to the validity of the results of the analysis.



# PERFORMANCE MEASURES, OBSERVER SELECTION, AND RECONNAISSANCE/STRIKE SYSTEMS EFFECTIVENESS

by

Herschel C. Self, PhD
Aerospace Medical Research Laboratory
Aerospace Medical Division
Air Force Systems Command
Wright-Patterson AFB, Ohio

OFFICE OF NAVAL RESEARCH
TARGET ACQUISITION SYMPOSIUM
NAVAL TRAINING CENTER, ORLANDO, FLORIDA / 14,15,16 NOVEMBER 1972

# PERFORMANCE MEASURES, OBSERVER SELECTION, AND RECONNAISSANCE/STRIKE SYSTEMS EFFECTIVENESS

by

Herschel C. Self, PhD
Aerospace Medical Research Laboratory
Aerospace Medical Division
Air Force Systems Command
Wright-Patterson AFB, Ohio

## ABSTRACT

Various measures of observer performance at the task of finding and recognizing targets are discussed along with consideration of the type of behavior required to do well on each measure. It is shown that to do well by different criteria requires different observer behavior. Experimental data is given which compares observers on different performance measures to illustrate the problem of selecting the most proficient observers. Conclusions are drawn about observer selection and its impact upon reconnaissance/strike systems effectiveness.

In a reconnaissance, strike, or reconnaissance/strike system there is a complex of equipment and men. It appears reasonable to assume that maximum effectiveness of the man-machine system would be obtained by utilizing the best men and the best machines. However, the best machines may be too expensive or unavailable for other reasons, and frequently one does not know who of a group of available men is the best. How does one select men and how does one select an affordable configuration of equipment? To what extent can quality of men make up for deficiencies in machines, and vice versa? The answers to such questions are extremely complex and are far from complete at the present time. However, it is certain that to arrive at answers one will have to deal with mission-relevant measures of human performance, equipment performance, and overall or system effectiveness.

Characteristics of the men, of the equipment, and of the situation or mission interact, so that variations in equipment and/or situation will have an effect on the men in the system. This is saying that a systems approach is essential when dealing with a system. Supermen and supermachines are futile unless they are matched to each others characteristics and to the mission, and there are no universal systems.

A corollary is that there are no universal men, that is men who are consistently the best by all performance criteria. It will become obvious why this is true when, later in this paper, mission-relevant performance criteria are examined. What are these criteria?

The first to come to mind is the number and percentage of targets that are detected, with the latter having more generality. An

ideal observer would detect, within the allowable time, all targets that were displayed with some minimum, at present unspecifiable, image quality. A real observer falls short of the ideal, often by a surprising or even shocking amount. Pressed for or limited by time, using improper search procedures, etc., he often misses targets displayed with excellent image quality and satisfactory image size. To obtain a high %D score, an observer may be forced to label many images of marginal resolution and/or of doubtful identity, "targets". He will make mistakes.

A measure of mistakes is "accuracy": The percentage of the objects that he designates as targets that truly are targets. An information-equivalent measure is percentage of false positives, i.e., objects mistaken for targets. An ideal observer would make no mistakes: His accuracy would be 100%. A real observer can't do this because real sensors in the real world don't display all targets with ideal contrast and resolution. Often some nontargets look more like real targets than do some targets. An observer who wants to obtain a high accuracy score has to be very sure of what he calls targets: He cannot call many marginal signatures targets without making errors. Thus, to score high on accuracy he has to be very cautious, while a high %D score is obtained by being very reckless.

In weapon delivery situations it is usually advantageous to detect and recognize targets while they are still at long slant ranges. Some weapon systems even require weapon release at long range. Since the slant range is usually decreasing rapidly due to motion of the aircraft, detection and recognition at long slant ranges is achieved by quick acquisition: Short reaction or detection times are required. However, at long ranges target images on the display are small and poorly resolved. To obtain a large target acquisition slant range score, i.e., quick detection, requires reckless behavior. If, to avoid mistakes, one waits until the displayed image is large and clearly resolved, then one will not obtain a "good" reaction time score. To obtain an excellent time score also requires ignoring targets first seen at close range and making snap judgments amounting to guesses on small target images at the top of the display. Thus, a good reaction time score is obtained at the cost of a low %D score and a low accuracy score.

Table 1 summarizes the behavior required for high scores on each of the three measures of performance. The table and the above discussion make it clear that one cannot do extremely well on all three performance measures because the types of behavior required to do so are incompatible or contradictory. The ideal observer would detect all of the targets (%D = 100%), mistake no nontargets for targets (accuracy = 100%), and would detect and recognize targets the instant that their images appeared on the display (reaction time = 0). A real observer cannot even come close to such behavior.

All three of the performance measures (%D, accuracy, and speed) are important. In a particular situation (or mission) the best men to use would be those who do best on the criteria of most importance in the particular situation. More accurately, men should be selected whose abilities on various measures of performance form the best match with the performance requirements of the task. For example, how well an observer does on the performance criteria depends upon both his ability and his emphasis. It is easy for observers to vary the relative emphasis that they give to different aspects of performance.

Thomas and Sadacca (1965)* demonstrated that image interpreters can vary their performance when the relative weight assigned to %D and accuracy is varied. The discussion in the present paper makes it apparent that to do well on one measure of performance may mean poor performance on another.

---

* Thomas, J.A., and Sadacca, R. Ability of Image Interpreters to Adapt Output to Varying Requirements for Completeness and Accuracy, USAPRO Technical Research Note 165, December 1965.

TABLE I.

| Performance Measure or Score | Behavior Required to Obtain a Very Good Score |
|---|---|
| %D | Label every possible image a target: Behave recklessly. Don't watch a possible target move down the display, growing in image size and resolution, before deciding, because this takes time that could be devoted to searching for other targets. Thus, quick reaction is important. |
| Accuracy | Be very cautious or conservative: Be sure it is a target before calling it one. Let its image move way down the display, becoming large and clear, before reacting: Take your time, better be slow than wrong. |
| Reaction Time or Slant Range | React Quickly: Don't wait for better resolution: Detect and recognize targets at the top of the display. Don't waste time searching the middle or bottom of the display, as targets acquired there are close and yield long reaction times. Reckless behavior pays off in long detection ranges. |

It is clear that measuring even highly trained men for ability so that the best man can be selected is not a simple matter. It would not be either simple or easy even if, for particular missions, one were given realistic "weights" for each of the three types of performance measures.

Despite the obvious difficulties in determining who is "best", it is important that it be done. This is true because, with either untrained or highly trained observers it is common to find some individuals who are far superior to others on any performance measure that one may select. Differences of 50% to 200% are common, and even larger differences are not rare. To obtain such large differences in the performance of man-machine systems by using superior sensors, for example, could require vastly more expensive equipment. For some state-of-the-art systems no amount of expended funds on research, development, and procurement could quickly (or even not so quickly) upgrade system performance as much as could more efficient observers.

Prior discussion has been on theory, but it is easy to illustrate some of the main points by examination of data from empirical studies. Three such studies by the author will be used. In the first study, fourteen observers had to find and recognize unbriefed targets on a rear-projection screen display of motion pictures which had been taken from an aircraft at a 340-foot altitude. The forward-looking camera had a field-of-view extending from a near slant range of 1950 feet to a far slant range of 4910 feet. At a simulated aircraft speed of only 60 knots they had to find targets of opportunity: men, cars, trucks, and heavy construction equipment. The mean scores for the observers and their rankings or relative positions on performance are given in Table II.

On %D, the best-to-worst score ratio was 76.54%/50.12%, or 1.53; the top observer on %D detected and recognized 53% more targets than did the bottom man. The ratio for accuracy was 91.25/46.49, or 1.96. On reaction time (detection + recognition) the average overall targets ranged from 8.62 seconds to 15.90 seconds, a ratio of 1.84. On screen position the ratio was 62.76/37.26,

or 1.68. Considerable range in scores is shown by the size of the standard deviations at the bottom of the table.

The rank or relative position of individual observers is of some interest since it permits a quick grasp of the relative standing of individual observers across two or more measures of performance. Table II shows that no observer was number one (best) on more than one measure and only one person, "M", was number 14 (worst) on as many as two measures. Looking at the last, or mean rank column, it can be seen that individuals "F" and "G" do relatively well on an overall basis. Thus, despite the incompatibility inherent in obtaining high scores on every performance measure, a few observers did relatively well overall.

TABLE II

OBSERVER PERFORMANCE WITH LOW ALTITUDE MOTION PICTURES SIMULATING AN AIRCRAFT SPEED OF 60 KNOTS

| | %D | | Accuracy | | Reaction Time | | Screen Position + | | Mean++ |
|---|---|---|---|---|---|---|---|---|---|
| Observer | Score | Rank | Score | Rank | Score | Rank | Score | Rank | Rank |
| A | 58.02 | 8 | 60.76 | 12 | 10.18 | 8 | 41.59 | 6 | 8.50 |
| B | 68.15 | 4 | 67.22 | 7 | 8.75 | 4 | 35.16 | 7 | 6.00 |
| C | 58.26 | 7 | 58.95 | 13 | 11.98 | 11 | 51.65 | 10 | 10.25 |
| D | 56.47 | 10 | 75.61 | 5 | 9.95 | 7 | 37.26 | 1 | 5.75 |
| E | 50.85 | 13 | 66.35 | 10 | 10.65 | 9 | 39.78 | 4 | 9.00 |
| F | 65.39 | 6 | 80.52 | 4 | 8.62 | 1 | 39.47 | 5 | 4.00 |
| G | 76.54 | 1 | 74.97 | 6 | 8.65 | 2.5 | 35.34 | 2 | 2.88 |
| H | 54.09 | 11 | 72.31 | 7 | 15.58 | 13 | 63.57 | 13 | 11.00 |
| I | 68.74 | 3 | 90.38 | 2 | 14.80 | 12 | 62.26 | 12 | 7.25 |
| J | 57.90 | 9 | 64.33 | 11 | 8.65 | 2.5 | 35.92 | 3 | 6.38 |
| K | 70.77 | 2 | 46.49 | 14 | 9.70 | 5 | 41.37 | 9 | 7.50 |
| L | 50.12 | 14 | 91.25 | 1 | 9.90 | 6 | 37.71 | 8 | 7.25 |
| M | 52.94 | 12 | 70.59 | 8 | 15.90 | 14 | 62.76 | 14 | 12.00 |
| N | 68.02 | 5 | 84.78 | 3 | 11.40 | 10 | 50.23 | 11 | 7.25 |
| Mean | 61.16 | | 71.75 | | 11.04 | | 45.29 | | |
| S.D. ** | 8.31 | | 12.49 | | 2.59 | | 10.73 | | |

The scores are the means of the separate scores on four types of targets: people, cars, trucks, and heavy construction equipment.

** Standard deviation of scores, a measure of scatter or individual differences.

+ Screen position is the distance down from the top of the display when the observer designated an object as a target.

++ Mean rank is the average of the four ranks.

The nature of the relationship between test scores on the various measures of performance becomes apparent when they are plotted against each other, two at a time, in scatter plots or diagrams. The six possible diagrams, with ranks plotted against each other, along with the rank correlation coefficient, for each set are given in Figure 1. Note that four of the six coefficients are less than .1, i.e., essentially zero: %D - accuracy, %D - screen position, accuracy-reaction time, and accuracy-screen position. Even the %D - reaction time correlation coefficient is a low .3462, which is significantly different from 0 ($t_{12} = 3.69$), but too small to be valuable in predicting one score from the other. The r' of .7525 for reaction time-screen position at detection is not very high. It is largely due to the bulk of the targets moving straight down the display at a uniform rate until detected so that for most targets, one measure is directly convertible to the other by a simple linear equation.

The story told by the values of the rank correlation coefficients is obvious from the scatterplots: instead of lying on a straight line or in a narrow ellipse, which would indicate a high degree of relationship, individuals are widely scattered. Only individuals "F" and "G" are in the lower left corner, indicating good scores, on all graphs. Only individuals H (5 out of 6 plots) and M (4 out of 6 plots) are fairly consistently in the upper right corner, indicating inferior performance.

A somewhat similar lack of all-around excellence in observer performance appears in a second study by the author. In this study twenty SAC radar navigators had to find unbriefed targets on a rear-projection screen display of side-looking radar imagery. An image of a strip of territory 41.5 nautical miles wide and 500 nautical miles long moved down the display at a rate which simulated an aircraft speed of 700 knots. Images required 215 seconds to move from the top of the display to the bottom, so that even at 700 knots observers were not rushed for time.

The basic performance data is given in Table III. Note that the percentage of targets detected varied from a low of 13.1% to a high of 38.5%, accuracy varied from 35.29% to 10.34%, and screen position varied from 1.9 to 7.2 clevenths of the way down the display. Thus, on any of the three performance measures there was a ratio of about three or greater between the highest and lowest scores .

The scatter plots of the ranks of the SAC observers are shown in Figure 2. Note that individual "C" is rather poor on all measures while observer "J" does not do badly on any measure. However, no person has a rank of lower (better) than 7 on all performance measures: No observer is superior on all three performance measures.

The last study used for illustrative data used low altitude strip photography displayed at slow helicopter speeds: 4.2 knots for one group of observers and 10.5 knots for another. Targets were manhole covers on streets and garbage cans in yards and alleys. Table IV gives the data %D was about 2:1 from best to worst observer at both of the speeds, accuracy in both was about 1.2 to 1, and the reaction time ratio was 25.3/5.8, or 4.4:1 at the slowest speed and 10.5/3.9 or 2.7:1 at the 10.5 knot speed. Thus, on %D and time (or target acquisition range), the range of individual differences was large, while all observers were rather close together on accuracy. Look at the six scatter plots in Figure 3, with particular attention to the lower left hand corners ("good" performance) and the upper right hand corners (bad performance). It may be noted that the same individuals don't occupy the same corners on all three plots at each speed. Few individuals are really close to the corner on even two plots. Thus, once again, it can be concluded that observers were not uniformly good or bad. Since none of the correlation coefficients were larger in magnitude than .5121, it follows that rank on any one performance measure has little predictive value for rank on another.

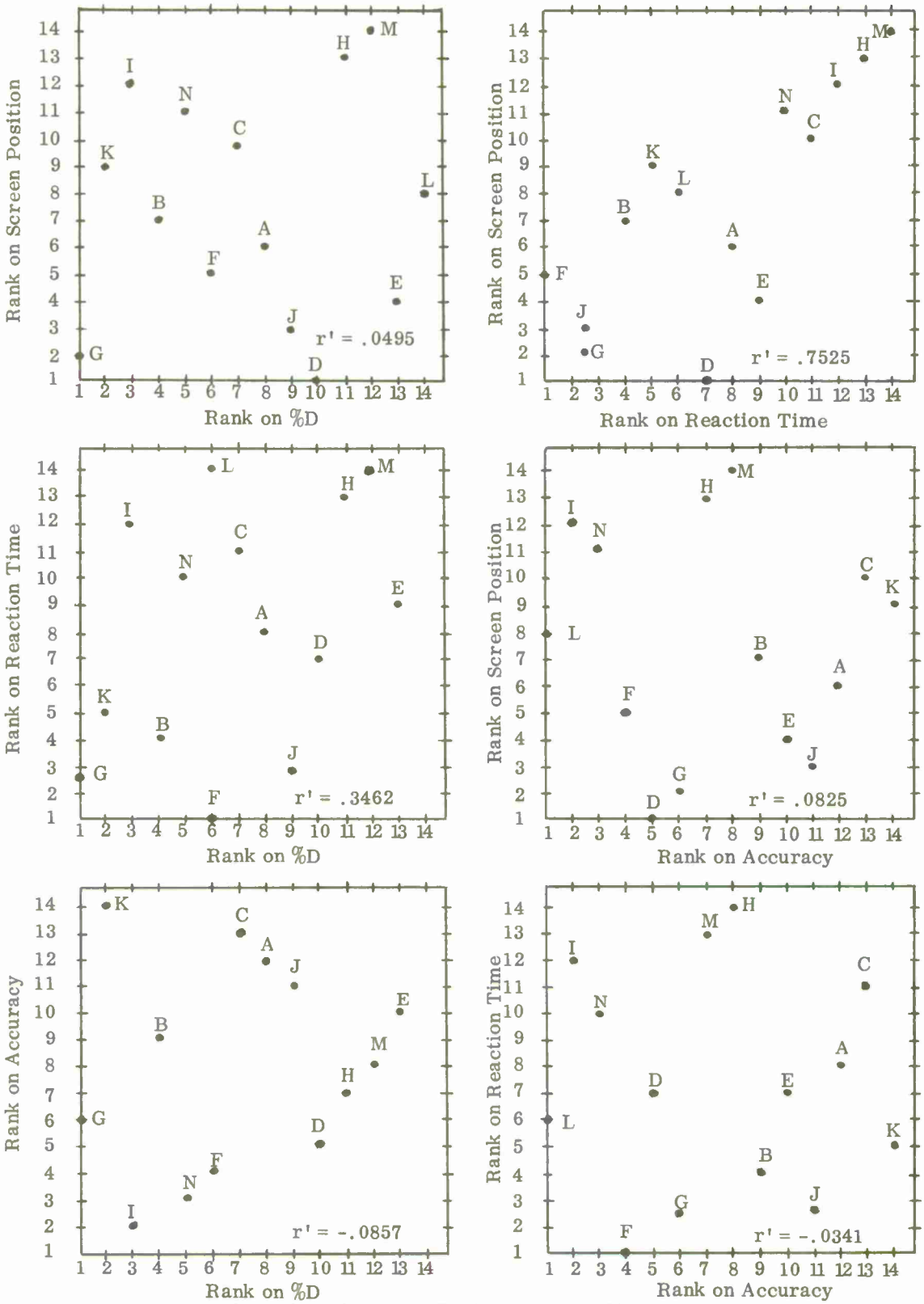



Figure 1. Scatter Plots of Performance Rankings of Ten Observers Using Motion Picture Imagery.

TABLE III

PERFORMANCE OF SAC RADAR NAVIGATORS WITH SLAR* IMAGERY
AT A SIMULATED AIRCRAFT SPEED OF 700 KNOTS

| Observer | Performance Measure | | | | | | Mean Rank |
|---|---|---|---|---|---|---|---|
| | %D | Rank | Accuracy+ | Rank | Screen Position++ | Rank | |
| A | 33.1 | 5 | 21.51 | 15 | 2.650 | 5 | 8.33 |
| B | 22.5 | 10 | 26.05 | 9 | 3.412 | 12 | 10.33 |
| C | 15.9 | 18 | 10.34 | 20 | 3.917 | 14 | 17.33 |
| D | 38.5 | 1 | 14.44 | 18 | 4.185 | 15 | 11.33 |
| E | 30.8 | 6 | 14.04 | 19 | 2.167 | 2 | 9.00 |
| F | 21.8 | 11 | 31.48 | 4 | 4.588 | 18 | 11.00 |
| G | 17.9 | 16.5 | 35.00 | 2 | 4.357 | 17 | 11.83 |
| H | 19.2 | 14.5 | 26.79 | 7 | 4.200 | 16 | 12.50 |
| I | 19.2 | 14.5 | 20.83 | 16 | 2.533 | 4 | 11.50 |
| J | 26.9 | 7 | 34.43 | 3 | 3.143 | 10 | 6.67 |
| K | 13.1 | 20 | 24.39 | 12 | 2.800 | 7 | 13.00 |
| L | 38.5 | 2 | 22.06 | 14 | 3.000 | 9 | 8.33 |
| M | 37.2 | 3.5 | 21.14 | 13 | 2.759 | 6 | 7.50 |
| N | 20.5 | 12.5 | 25.00 | 11 | 7.250 | 20 | 14.50 |
| O | 37.2 | 3.5 | 20.42 | 17 | 1.923 | 1 | 7.17 |
| P | 24.7 | 8 | 26.76 | 8 | 2.316 | 3 | 6.33 |
| Q | 20.5 | 12.5 | 31.37 | 5 | 2.875 | 8 | 8.50 |
| R | 15.4 | 19 | 27.91 | 6 | 3.167 | 11 | 12.00 |
| S | 23.1 | 9 | 35.29 | 1 | 3.667 | 13 | 7.67 |
| T | 17.9 | 16.5 | 25.45 | 10 | 6.571 | 19 | 15.17 |
| Mean | 24.70 | | 24.74 | 10.50 | 3.574 | 10.50 | 10.50 |
| S.D. | 8.30 | | 6.93 | 5.95 | 1.371 | 5.95 | 3.05 |

* SLAR = Side-Looking Airborne Radar.

\+ Accuracy = (100) X (Number of Targets Detected)/(Detections + False Positives) = (100) X (Number Correct)/(Number of Responses) = %Correct.

++ Screen position in elevenths of the way down the 14-inch display. Since 215 seconds was required to move from top to bottom, each eleventh represents 19.5 seconds.

NOTE: Targets were unbriefed dams, railroad yards, industry, tank farms and airfields.

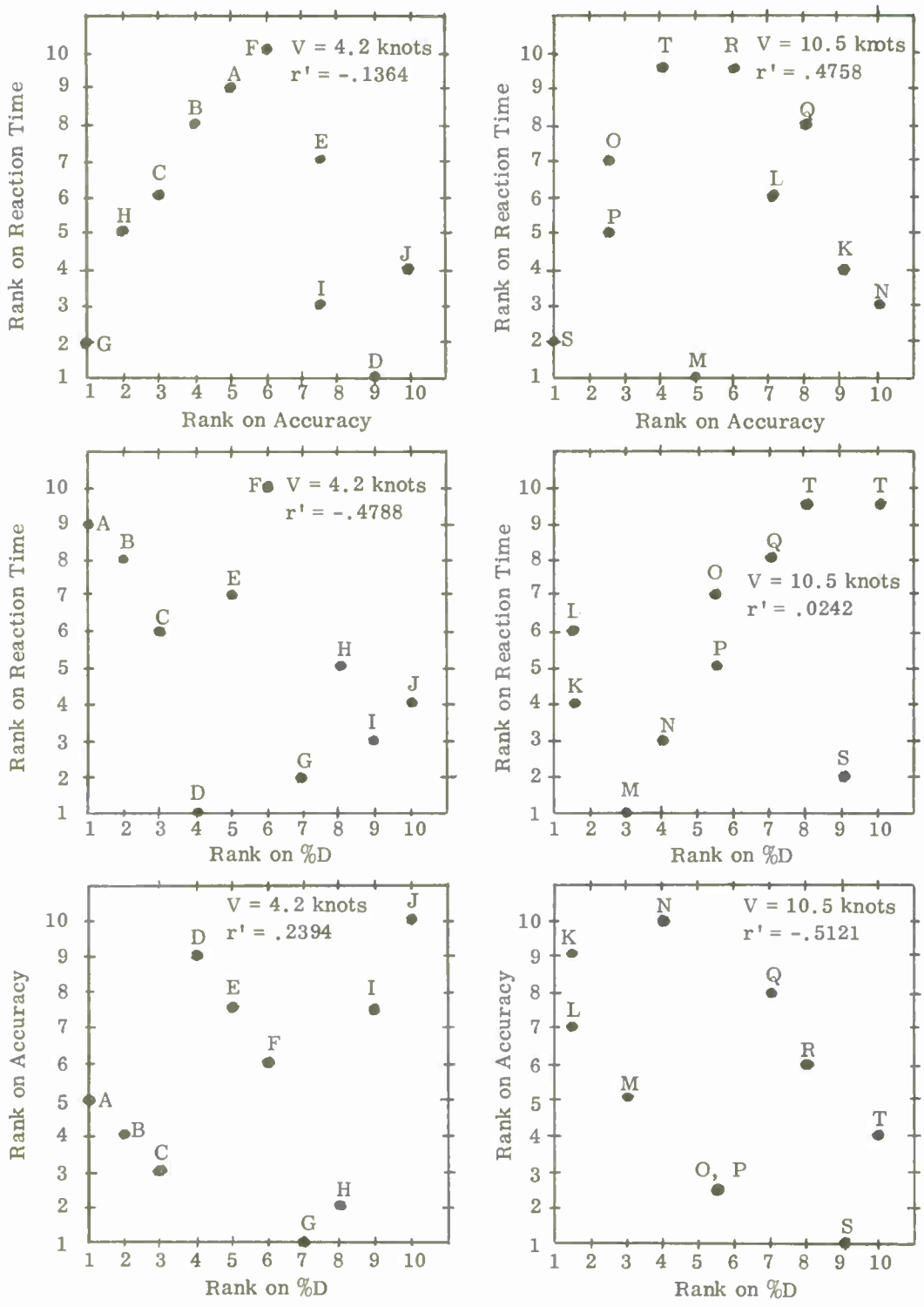
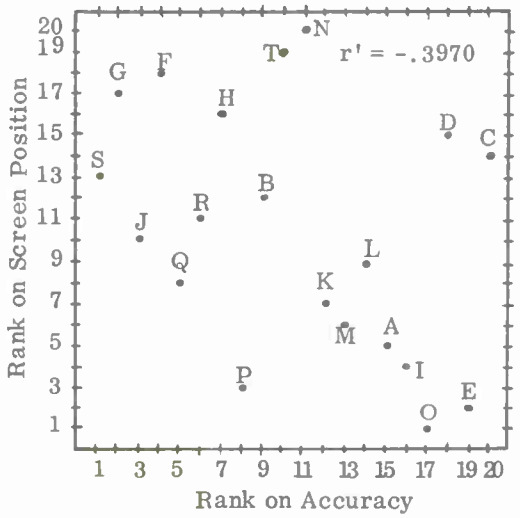
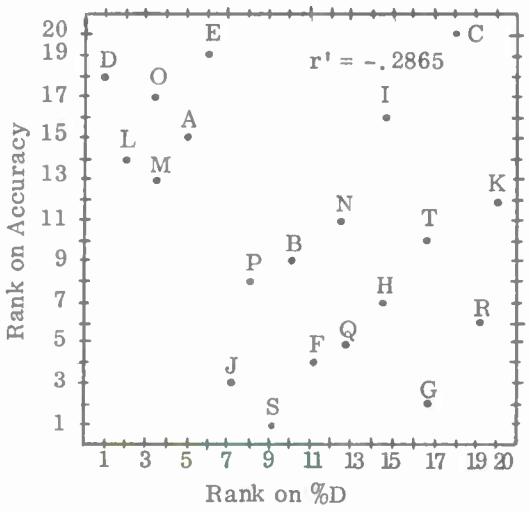



Figure 3. Scatter Plots for Ranks on Performance Measures for Observers Using Strip Photography.



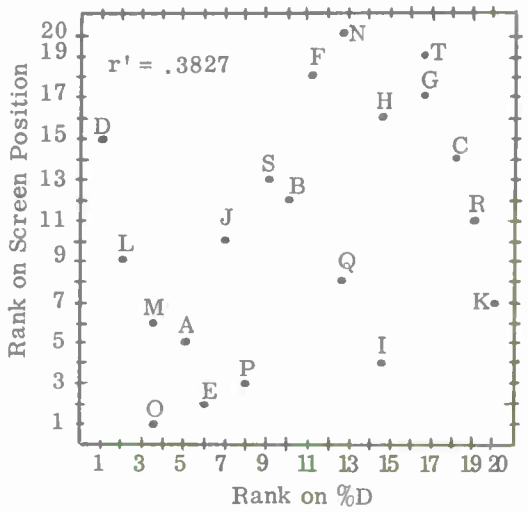




Figure 2. Scatter Plots of Performance Rankings of Twenty SAC Radar Navigators Using Side-Looking Airborne Radar.

TABLE IV

OBSERVER PERFORMANCE WITH LOW ALTITUDE STRIP PHOTOGRAPHY
SIMULATING AIRCRAFT SPEEDS OF 4.2 AND 10.5 KNOTS

| | SIMULATED HELICOPTER SPEED | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 4.2 KNOTS | | | | | | 10.5 KNOTS | | | | | |
| Observer | %D | Rank | Accuracy | Rank | Reaction Time | Rank | Observer | %D | Rank | Accuracy | Rank | Reaction Time | Rank |
| A | 39.22 | 1 | 88.89 | 5 | 19.0 | 9 | K | 24.02 | 1.5 | 84.48 | 9 | 7.4 | 4 |
| B | 37.25 | 2 | 89.41 | 4 | 17.6 | 8 | L | 24.02 | 1.5 | 85.96 | 7 | 7.7 | 6 |
| C | 33.83 | 3 | 87.34 | 3 | 11.0 | 6 | M | 23.53 | 3 | 88.89 | 5 | 3.9 | 1 |
| D | 33.33 | 4 | 73.91 | 9 | 5.8 | 1 | N | 21.08 | 4 | 81.13 | 10 | 7.1 | 3 |
| E | 31.86 | 5 | 83.33 | 7.5 | 13.1 | 7 | O | 20.10 | 5.5 | 97.62 | 2.5 | 7.9 | 7 |
| F | 27.45 | 6 | 83.58 | 6 | 25.3 | 10 | P | 20.10 | 5.5 | 97.62 | 2.5 | 7.5 | 5 |
| G | 20.59 | 7 | 91.30 | 1 | 8.1 | 2 | Q | 17.16 | 7 | 85.37 | 8 | 9.3 | 8 |
| H | 19.61 | 8 | 90.91 | 2 | 9.9 | 5 | R | 14.71 | 8 | 88.24 | 6 | 10.5 | 9.5 |
| I | 19.61 | 9 | 83.33 | 7.5 | 8.7 | 3 | S | 13.73 | 9 | 100.00 | 1 | 5.0 | 2 |
| J | 17.65 | 10 | 67.92 | 10 | 9.2 | 4 | T | 13.73 | 10 | 96.55 | 4 | 10.5 | 9.5 |
| Mean | 28.04 | | 83.99 | | 12.77 | | | 19.22 | | 90.59 | | 7.68 | |
| S.D. | 8.11 | | 7.65 | | 6.07 | | | 4.15 | | 6.72 | | 2.15 | |

NOTE (1): Observers were arranged in order on %D scores before being assigned letter designations in the above table.

NOTE (2): Targets were unbriefed manhole covers in streets and garbage cans.

## SUMMARY AND CONCLUSIONS

1. The type of behavior required for doing well according to one performance measure may be incompatible with that required for doing well on another.

2. Examination of data from human factors studies reveal that observers differ widely on performance scores, a range of 50% to 300% not being unusual on any one type of measurement.

3. There are very few, if any, observers who rank very high (or very low) on all three of the most obvious performance measures: percent of targets detected, accuracy, and speed (time or slant range or screen position at detection). A few individuals do well on two of the three measures, but not extremely well on both.

4. The correlation coefficients between scores on different combinations of the three main criteria are usually too low to be of practical use: Prediction of scores on one measure from scores on another is practically not possible.

5. These facts indicate that the observer selection problem is important but very difficult. Theory and data both indicate that one must select for particular missions or purposes, rather than for general purposes: Few, if any, observers are "tops" in everything. Effective selection devices are needed.

6. More research is needed, particularly on the absolute and relative scores of the same observers in different situations and with instructions giving different emphasis on the relative importance of %D, accuracy, and acquisition range.

The research reported in this paper was conducted by personnel of the Aerospace Medical Research Laboratory, Aerospace Medical Division, Air Force Systems Command, Wright-Patterson Air Force Base, Ohio. This paper has been identified by Aerospace Medical Research Laboratory as AMRL-TR-72-86. 

# THE MYTHOLOGY OF TARGET ACQUISITION SYSTEM DESIGN AND PERFORMANCE

by

Lucien M. Biberman
Institute for Defense Analyses
Arlington, Virginia

OFFICE OF NAVAL RESEARCH
TARGET ACQUISITION SYMPOSIUM
NAVAL TRAINING CENTER, ORLANDO, FLORIDA / 14,15,16 NOVEMBER 1972

# THE MYTHOLOGY OF TARGET ACQUISITION SYSTEM DESIGN AND PERFORMANCE

Lucien M. Biberman
Institute for Defense Analyses
Arlington, Virginia

## Abstract

This paper considers the factors long associated with human performance and vision and contrasts such factors with those that have more recently been found to have greater relevance. In particular, a brief review is made of past experiments and present programs and compares and contrasts the important factors emerging from each. A series of examples are given in both visual and acoustical analogs to point up some of the more commonly believed mythology.

This afternoon I shall concern myself only with those factors that concern transfer of information from an image to an observer, and shall concentrate on the mythology surrounding television-type imagery.

All that glitters is not gold; the hand is quicker than the eye; and the "sharp" television tube often interferes with the viewer's seeing the picture that is really there. But let's get to that a little later.

The sudden interest in low light level technology in the early sixties and particularly in the first of the intensifier-image orthicon television cameras, resulted in a flurry of field evaluations which by and large poorly displayed various pictures of various "military targets" under various circumstances with little or no control of conditions.

Yet some good work was done in which the same levels of "resolution" as were expected from low light level television were simulated by projected photographs, but raster effects and noise characteristics were not understood and thus ignored or were badly simulated. As a result no real relationship was established that would predict, based upon laboratory trials or photographic simulations, how well a man actually using a television set of a given type would perform. At this point, I am quite careful not to say anything about the relationship between system quality and observer performance. The reason is that the television systems were specified, ordered, designed and delivered on the basis of a number of system parameters that had little or nothing to do with image quality or visual performance, i.e., the basic factors that allowed people to see things on a television display.

There were some who attempted to carry out human factors experiments with television earlier, but some of the first quantitative work of interest to me was carried out by Ronald Erickson and his associates under conditions in which television images of relatively high quality were used to test human ability to perform visual tasks.

Dorothy Johnson (1968) conducted

a series of experiments concerned with target recognition as a function of slant range, horizontal resolution and shades of gray. The measured performance parameters were target recognition time and probability. Ms. Johnson's results showed correlation between the parameters studied.

Scott, Hollanda and Harabedian (1970) carried out a series of simulations of TV imagery through the use of Frank Scott's line scan generator. This work was like that of Hemingway and Erickson on symbols except vehicular images were used as inputs and perspective was an important but not consistent parameter. They reported the identification and classification accuracy of noiseless, static line-scan images as a function of the number of scans per vehicle (4, 6, 9, 13.5, 20 and 30 scans) and angle of view (nadir and 45 degree obliquity angle).

In a related work, Gaven and Tavitian (1970) study the effect of shades of gray and the number of scan lines on observer performance. They studied the informative value of sampled imagery as a function of the number of sampling elements and the number of bits used to encode the gray levels in the imagery.

The above series of experiments are not the only work that has been done, but are representative of a class of psychophysical experiment that relates performance to display parameters that may indeed not be the independent variable.

The line scan generator of Scott (1967) produced a series of line scanned imagery typically illustrated by the few samples in Figs. 1 and 2. It is obvious that there are other more basic factors that drive the quality of line scan imagery which produce the dependent variables such as "lines" and "gray scale."

The raster lines also produce false imagery caused by the interaction of the signal inherent in the raster line frequency with frequencies present in the imagery, i.e., sum and difference effects inherent in any sampling system.

Though the effect of the sampling problem has been explained--first in the middle thirties and again in the fifties--it is coming up again in the seventies as serveral whole new families of flat displays or picture tubes emerge. These new displays like those in small computers, cash registers and the like are composed of a large number of individual elements that glow more or less brightly to form a picture. They are displays made of many discrete elements and differ from usual images in the way that newsprint pictures differ from good photographs. That is, the images are sampled rather than continuous and sampling can produce lots of problems whether the sampling is only in one direction like the lines in a television picture or in two dimensions like the dots in a newsprint picture. The more distinct and sharp these lines or dots are, the more trouble they cause. Also the farther apart or the coarser the pattern, the worse the troubles. This is obvious, see figures on raster and Moire effects, Figs. 3 and 4, but it is really a small part of the whole story.

These factors were predicted before the birth of television as we now know it. The difficulties due to sampling in a series of discontinuous lines were clearly set forth.

Mertz and Gray in 1934 predicted the raster effect on image quality. Twenty years later, after patenting means to remove some of the injurious effects predicted by Mertz and Gray, Otto Schade (1953) published a complete quantitative treatment supplementing the paper of Mertz and Gray.

Four years after Schade's paper, F. T. Thompson (1957) reported on personal preferences in viewing television imagery and the effect that removing raster structure had upon these preferences.

Thompson describes a variety of means for accomplishing a reduction in the appearance of individual lines. He conducted a series of experiments with various raster frequencies. The results strongly confirm each other and thus we quote only his results.

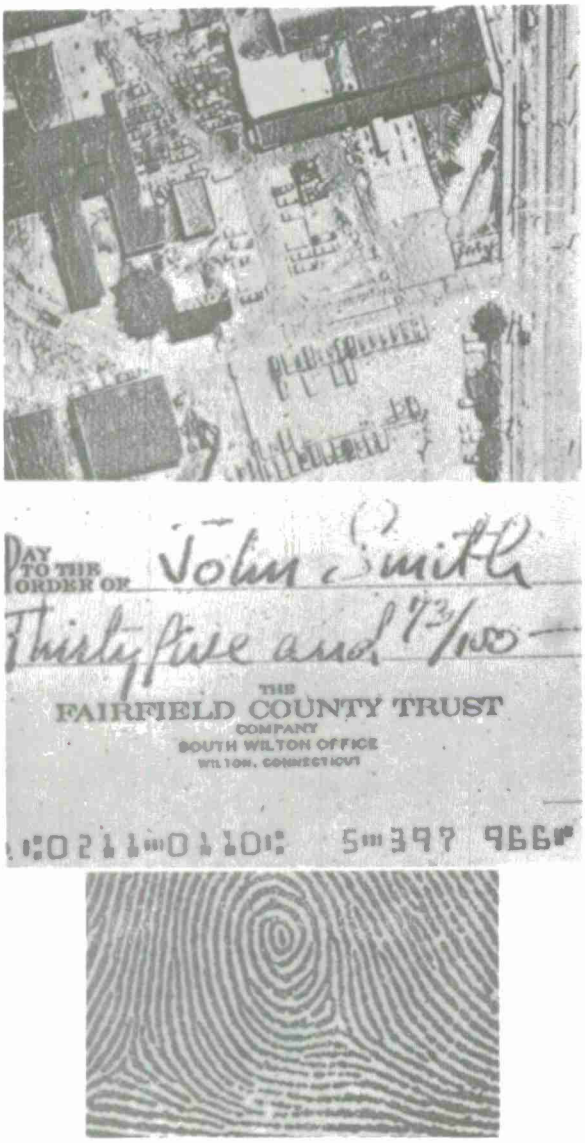



Figure 1. Line-Scan Image Made With LSIG in Electronic Mode Illustrating Low Spatial-Frequency Attenuation (1.50 scans per mm)

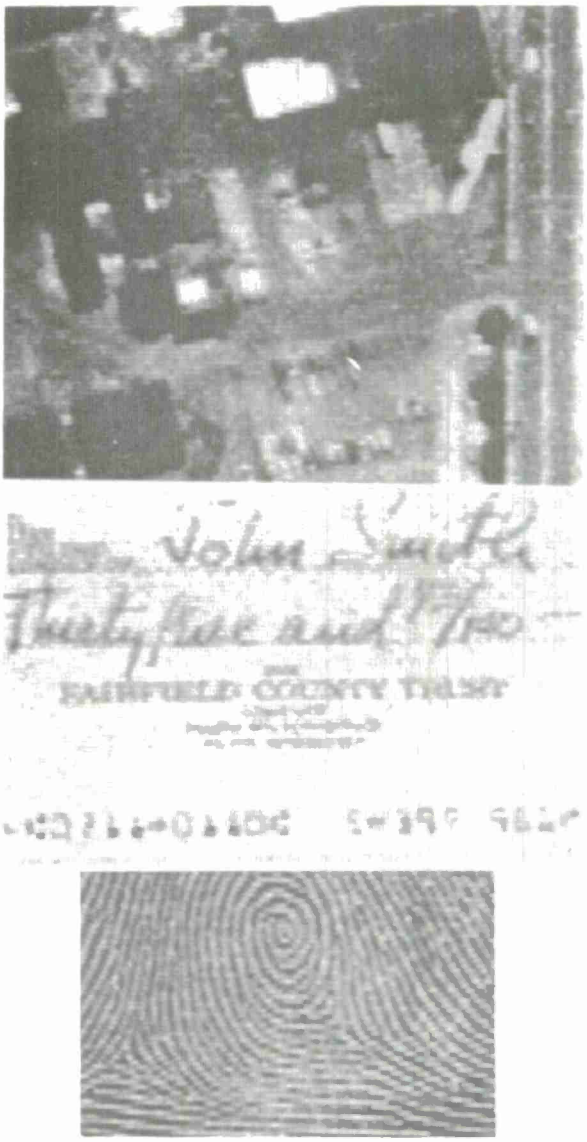


Figure 2. Line-Scan Image Made With LSIG in Electronic Mode – Gaussian Noise Added (1.5 scans per mm)

Figure 3a. This Pattern as Object, Sampled Through

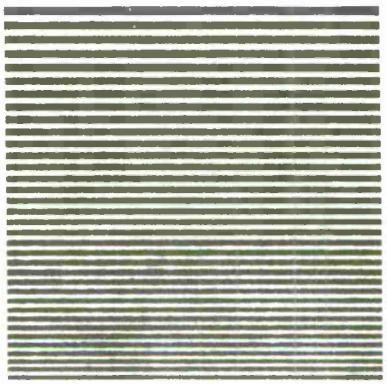


Figure 3b. This Pattern as a Raster,

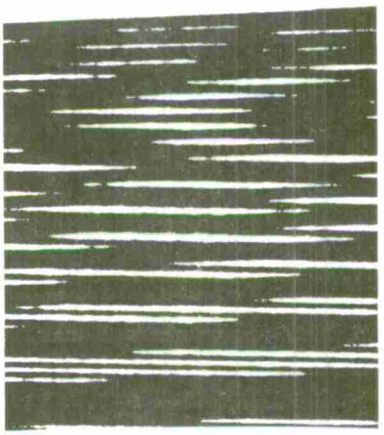


Figure 3c. Yields This Image

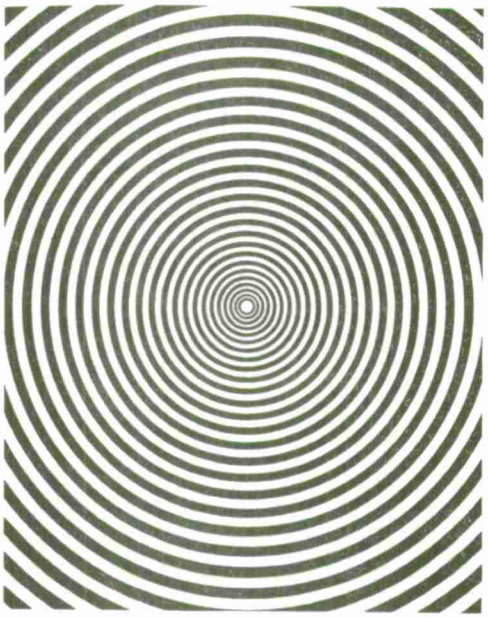


Figure 4a. A Concentric Pattern for Moire Experiments

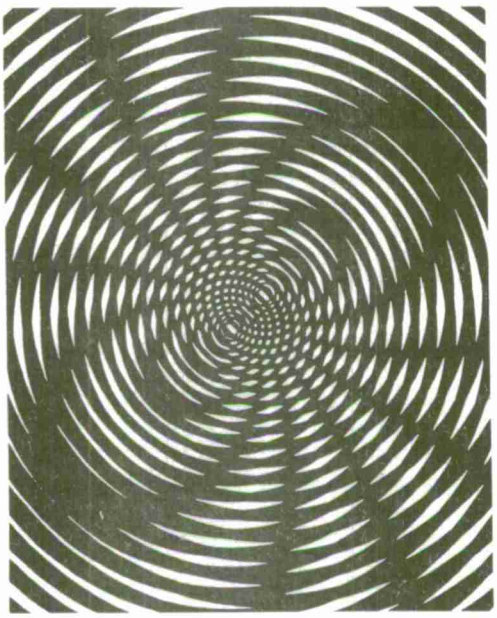


Figure 4b. A Moire Pattern Caused by a Small Displacement of the Previous Figure

False Images Resulting from Aliasing

"It would be desirable to suppress or eliminate the line structure in television pictures thereby allowing the larger receivers to be comfortably viewed at reduced distances."

Thompson thus indicated that the viewers moved away from the television set until they could no longer see the raster lines (or the equivalent fine image structure along a line). When the raster was removed, the viewers moved back to the same shorter distance they chose to observe raster-free imagery and thus could observe greater detail. These effects are shown in Figs. 5 and 6 from Thompson's paper.

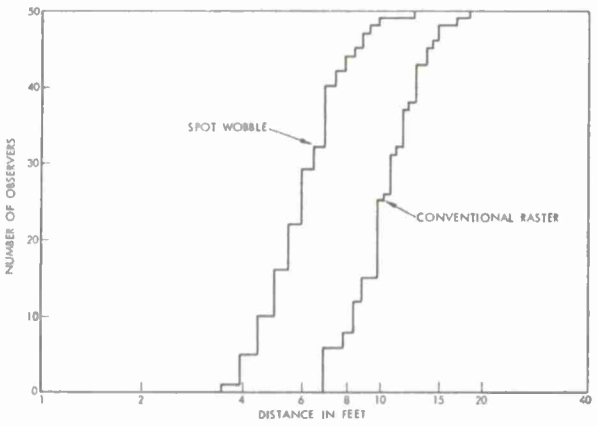



Figure 5. Observers Sitting Closer Than a Given Distance (525 Line Raster on 24 Inch Receiver)

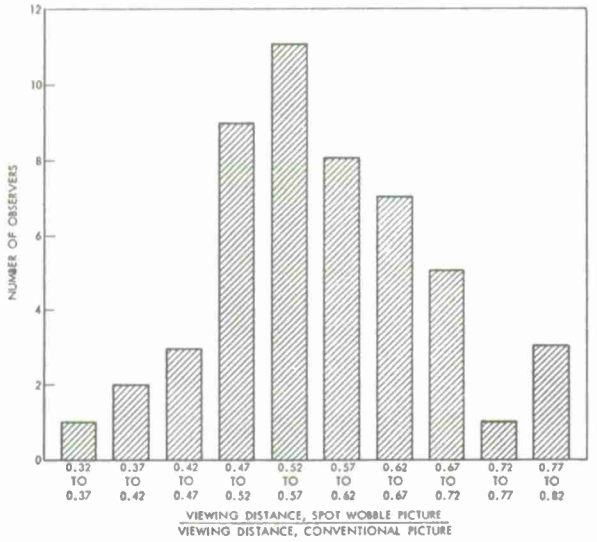



Figure 6. Effect on Spot Wobble on Viewing Distance (525 Line Raster)

Fifteen years after Thompson's experiment and eighteen years after Schade's SMPTE paper, there still is a determined effort to get more information to pilots in aircraft through television or similar image-forming equipment. Though Schade has done the calculations and Thompson has done an irrefutable experiment, system designers still strive for greater amounts of "limiting resolution" in camera tubes and ignore the effect of the raster in the camera tube and especially in the display.

In 1969, in search of a new domestic television set, I read ratings, talked to knowledgeable colleagues and went out to shop for a color television receiver. In the course of visiting a number of stores, I was strongly impressed by the actions and apparent criteria of many other shoppers. A surprisingly large number of prospective buyers got up close to the picture tube and examined the raster structure. Most seemed very pleased with pictures in which the "lines" were sharp and clean. I intruded to ask a number of buyers so engaged how they were making their choice since they obviously were making a close technical (if subjective) comparison. Flattered at my recognition of them as men of discrimination and sound judgement, they invariably pointed out to me that they were looking for a set that was in good adjustment and sharp focus--"just look at how sharp the lines are,--real good detail in that set."

Now let me show you how silly this criterion really is. Figure 7 shows a picture of a map as seen on a two dimensional matrix display. The whole picture is quite good. It's just hard to read detail.


Figure 7. Map – Two Dimensional Matrix Display

The second figure is a good magnified picture of a small section. There is no more detail because the raster gets in the way. If I defocus the picture and let everything get blurry, the fine raster lines get blurry faster than the larger type and we can read more than we previously could with a sharp picture. (See Figs. 8 and 9.)

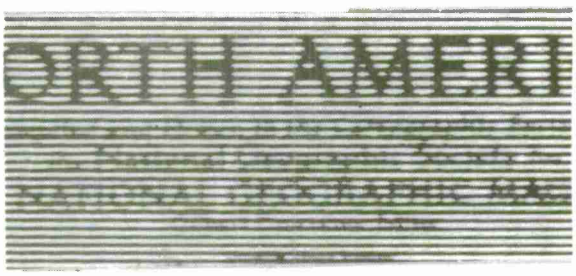
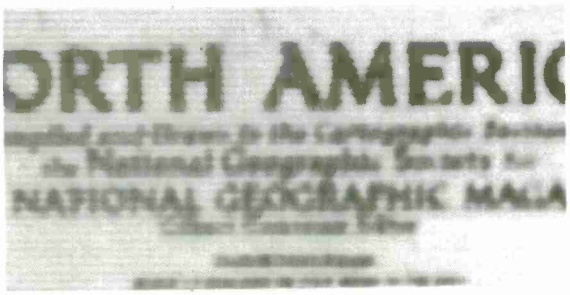


Figure 8a. One Dimensional Raster Imagery Buried in Raster


Figure 8b. Two Dimensional Raster Imagery Buried in Raster



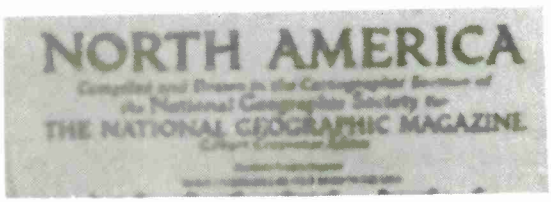



Figure 9. Defocussing Blurs Raster Faster Than Type

Defocussing loses information, but loses it faster at high frequencies; thus, defocussing loses raster before it loses the larger type on the maps.

Well-designed displays should never have a raster pattern that is visible from a normal viewing distance. The displays we get will continue to be poor as long as we don't understand these facts and are willing to buy television sets that mutilate imagery with sharp black and white lines. One must be sure that the set has fuzzy lines that are almost invisible but a sharp picture, NOT vice versa.

At the risk of becoming somewhat less than accurate let me explain that the effect of the sharp lines in a picture are somewhat like a piercing whistle that bothers us when we are trying to listen to code on a radio.

Remove the piercing scream of the whistle and the soft code can be heard distinctly. In the same way remove the "loud" black and white lines, or make them softer to be able to "hear" the "softer" images on the tube.

As the new displays come into the development cycle it is most important that the line structure be minimized in order to maximize the usefulness of the new displays to the human viewer.

We have said necessary but nasty things about lines in sampled imagery. We feel that the previous experiments overlooked the fact that the raster lines need not, and should not, be visible.

Thus, though we do not disagree with the conclusions drawn from the psychophysical television experiments previously noted, we would like to point out that by changing the properties of the spot sizes of "apertures" used with a display raster, but not the raster frequency or basic form, the results would be strongly affected.

Actually, one must conclude that the angular subtense and number of raster lines do affect the legibility of symbols, however, other primary, more basic factors of the raster can have equal or greater effect and thus the raster parameters must be specified.

Now, let me show a few slides and perform a few fun experiments with you.

First let me show the statistical aspects of video communications in a slide made, I believe, by George Morton for his magnificent article of Applied Optics a few years back.

Figure 10 shows the resulting effect on information content in an image caused by limiting the number of bits of light or photons making up the image. It's a beautiful example of Poisson statistics.

Next let me show you the effect of too much of a good thing--more light--and what modulation in imagery means to you.

Look at the next series of figures, Fig. 11, same as before but now let me throw some white noise on the subject. Please note I have not removed the information present in the figures--it's still all there. It's just that you with your people-sized computational center between your eyes can no longer meaningfully process in light patterns when the random variations inherent in the high level of flux from our second source equal or exceed those variations that convey information to us. You could do the job but only with much more processing or longer integration times.

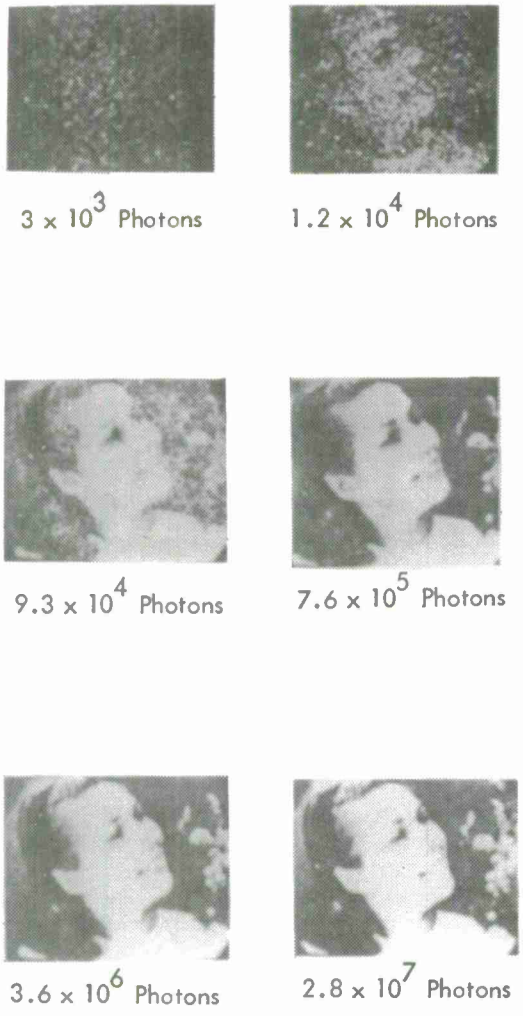



Figure 10. Images Formed by Scintillations - Series of Photographs Showing the Quality of Pictures Obtained with Various Numbers of Photons (or photoelectrons when the quantum efficiency is less than unity)

Figure 11. Effect of Adding Noise (more unfocussed light) to Imagery Shown on Left

It is a little like listening to an itinerant violinist on a subway platform as the express passes. I'm sure that the violinist's talent did not degrade as the train rushed by. Let's listen to that analog. In vision, it's like my imagery now on the screen.

The problem can be and is much more selective and should be clearly understood, but I very much doubt if anyone here ever took the trouble to remedy the situation.

I'm talking about the non-linear effects that occur when one acoustic frequency greatly exceeds in amplitude a reasonably good signal, perhaps a couple of octaves removed.

Like what? Like a 5KC hiss in a super-regenerative receiver saturating your neural networks so that you can't read the coded signals. How do you clean that up? Filters, of course. In acoustics it's like this. (Audio demonstration.) How in Optics? See Figure 12.

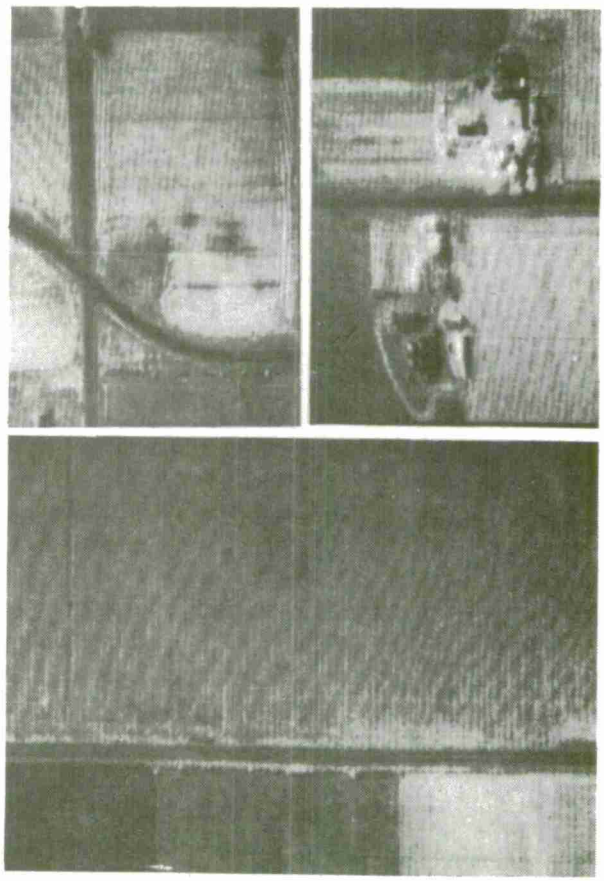


Figure 12. Moire Patterns From Plowed Fields Due to Sampling

The fascinating work that is going on now in careful psychophysical experiments performed by a combination of people who understand both physics and good experimental design practice in psychology are yielding good hard meaningful results.

Unfortunately in the recent past too much work has been done correlating human visual performance primarily with things like the number of raster lines or shades of gray in television imagery. I hope my earlier demonstrations convinced you that nothing is important about either one unless one understands the signal level inherent in the raster and in the image to be viewed in spite of that raster. The shades of gray likewise are directly related to the dynamic range of signal to noise ratio.

In the experimental material reviewed earlier the relationship of the viewer and his ability to accomplish certain visual tasks was consistently related to a number of parameters primarily the number of scan lines and the number of shades of gray in the image.

Let me repeat, though we do not disagree with the conclusions drawn from these particular experiments we would like to point out that by changing the signal-to-noise ratio and/or the properties of the spot sizes or "apertures" used with a display raster, but not the raster frequency or basic form, the results would be strongly affected. Actually one must conclude that the angular subtense and number of raster lines do affect the legibility of symbols, however, other basic factors of the raster can have equal or greater effect and thus the raster parameters must be specified.

For the particular conditions of the experiment, the MTF of the camera tube, the effective prefiltering of the lens used, etc., result in a level of aliased signal that causes definite and perceptible deterioration in image quality of the type that Schade shows in his examples, shown in Figure 13.

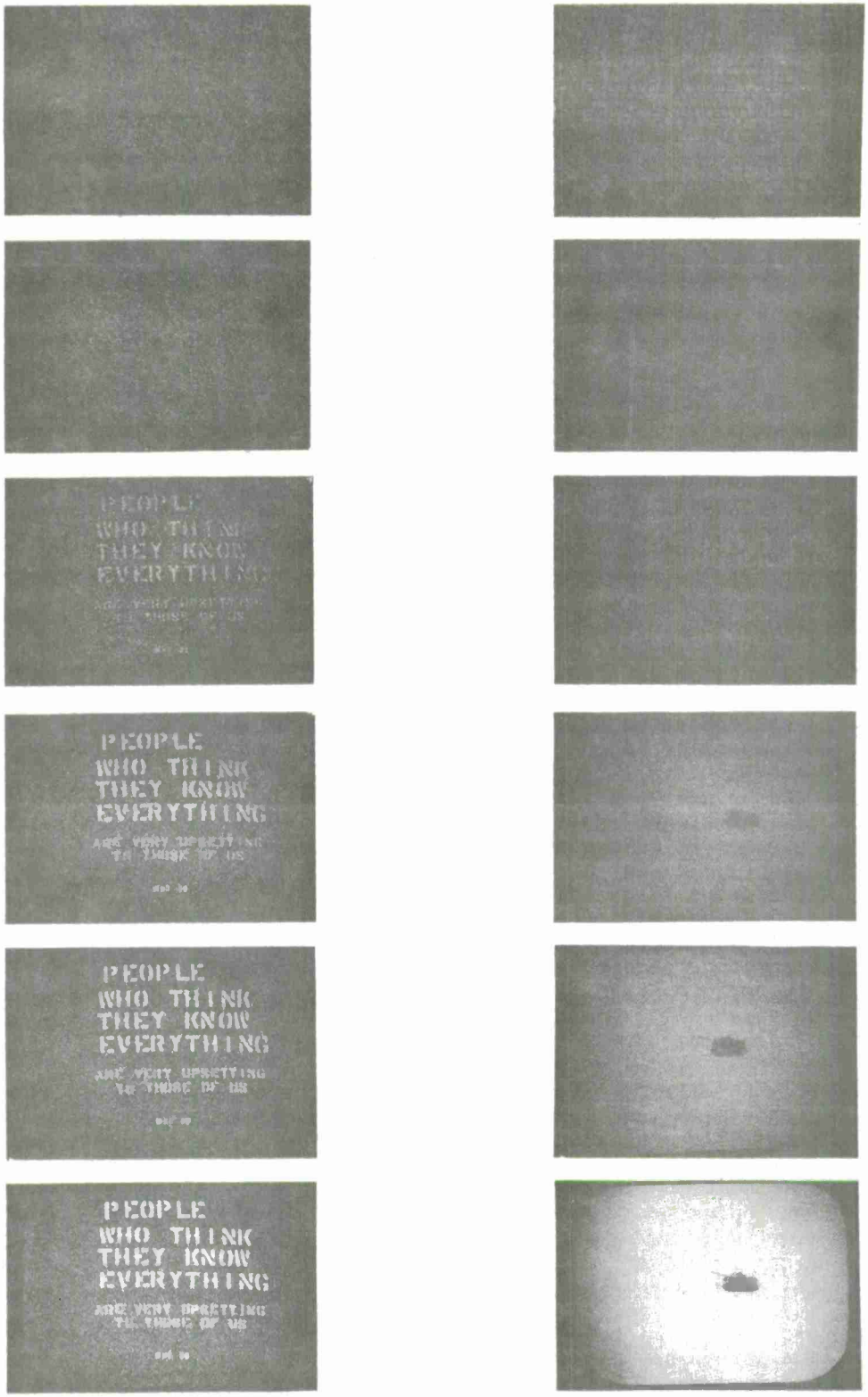



Figure 13. Two Examples of Image Quality as a Function of Signal-to-Noise Ratio for a Fixed Line Number Display

It therefore seems a fruitless task to relate visual capability to raster lines in general when such an effect is highly related to the design and adjustment of an individual specific television system. It is clear that the raster visibility will be an uncertain, variable, and real problem as long as there is little or no concern about its existence or degree of interference--and presently there is almost no such concern, since by and large the buyers of television sets designed for domestic entertainment not only accept as fundamental the presence of the raster lines but often choose a set on the supposition that a clear, sharp, highly visible raster is an indication of a well-designed set in good adjustment.

Much thought needs to be given about the real findings of the experiments discussed earlier. As is shown by Schnitzler and Rosell, modulation and/or signal-to-noise ratio can be used to indicate the limitations in both "resolution" and "shades of gray" in imagery.

It is clear that the recognition process and related visual processes are related to "resolution" and "shades of gray" but it is equally clear that both of these are dependent upon the signal-to-noise ratio in the imagery as a function of spatial frequency, the latter determining the former.

It is interesting to note that shades of gray represent steps of $\sqrt{2}$ in luminance of the image or resolution element. Obviously with a limited signal-to-noise ratio the number of shades of gray are automatically limited; two shades of gray calling for a dynamic range in S/N of $(\sqrt{2})^2$ or 2:1 while eight shades of gray call for $(\sqrt{2})^8$ or a signal-to-noise dynamic range of 16:1.

The interaction between object size and shades of gray is highly predictable. Since the response of a television system decreases in a predictable manner as image sizes become smaller, signal-to-noise ratios decrease as images become smaller until finally the signal-to-noise ratio drops to a just detectable level when the size of the target reaches the "limiting resolution" of the system. This point is a function of the radiance of the object and its contrast against the background and the characteristics of the TV equipment. At limiting resolution there normally is no gray scale. There is less than one shade of gray difference between target and background, for anything smaller, the signal-to-noise ratio drops below a detectable SNR or, as more usually expressed, below a discernible contrast level at the given level of display brightness.

If the properties of the experimental equipment were known, the results obtained should be predictable from these parameters as well as from the derived parameters of resolution and shades of gray, which are subjective manifestations of signal to noise.

As stated before, a TV system with the lens capped, produces the same display raster, with the same bright and dark horizontal stripes as the raster associated with a bright and clear image. The first has a good signal-to-noise ratio but only at the raster frequency and then only in the vertical direction and thus has no informative content.

The experimental and theoretical work of Rosell and the theoretical background derived from deVries, Rose, Schade, Coltman and others do not directly of themselves yield an estimate of the performance of a man in, for example, distinguishing between the body styles of various automobiles. Yet that same set of theories coupled with a knowledge of the equipment characteristics and the scene characteristics allow one to predict the signal-to-noise ratio vs frequency on the display of the given equipment, the raster parameters and therefore the number of lines on a given image, and the shades of gray in that image.

If the latter two determine the informative content of imagery, the combination of a knowledge of the scene and the equipment should allow

one to predict the level of achievement of the given visual task which was predictable from the subjective characteristics. Perhaps more to the point knowing the characteristics of the scene and the observers' requirements, one should be able to specify the level of system specification necessary to achieve a given visual task.

Actually the computational process for determining probability of detection or identification, or the inverse problems of determining system parameters necessary to achieve the required probability is considerably simpler than it sounds in the above discussion. The process and methods of calculation are presented in a straightforward manner towards the end of Rosell's Chapter 5 in "Perception of Displayed Information" by Biberman et al.

If there are those who still believe that raster lines really do control an observer's ability let me show a series of slides in which everything is constant except the signal-to-noise ratio.

Finally let me show you a great picture where the number of lines across a subject meets all requirements but the signal to noise is zero.

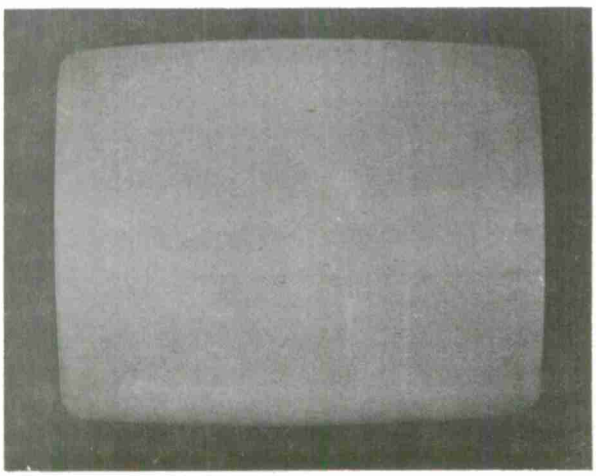


Figure 14. Picture of a Tank at 1000 Meters, 120 TV Lines per Tank Dimension, Signal-to-Noise Equal to Zero

BIBLIOGRAPHY


Beurle, R. L. and Hills, B. L., "Visual Perception with Electronic Imaging Systems," Report No. J.5, 1968, University of Nottingbean, Department of Electrical Engineering, March, 1968.

Beurle, R. L. and Daniels, M.V., "Visual Perception with Electronic Imaging Systems," Report No. J.8, 1968, University of Nottingbean, Department of Electrical Engineering.

Biberman, L. M. et al, "Perception of Displayed Information," Plenum Press, in preparation due Spring 1973.

Coltman, J. W. and Anderson, A. E., "Noise Limitations to Resolving Power in Electronics Imaging," Westinghouse Research Labs., Pittsburgh, Pa., Proceedings of the IRE, May 1960, p. 858.

Erickson, R. A., Linton, P. M., Hemingway, J. C., "Human Factors Experiments with Television," NWC, China Lake, California, Report NWC TP4573, October 1968.

Erickson, R. A. and Hemingway, J. C., "Visibility of Raster Lines in a Television Display," JOSA Vol. 60, No. 5, pp. 700-701, May 1970.

Fellgett, P. B. and Linfoot, E. H., "On the Assessment of Optical Images," Vol. 247, A. 931, p. 46, Feb. 17, 1955.

Gaven, J. V., Tavitian, J., Harabedian, A., "The Informative Value of Sampled Images as a Function of the Number of Gray Levels used in Encoding the Images," Photographic Science and Eng., Vol. 14, No. 1, pp. 16-20, January 1970.

Hemingway, J. C. and Erickson, R. A., "Relative Effects of Raster Scan Lines and Image Subtense on Symbol Legibility on Television," Human Factors, Vol. II, No. 4, p. 337-338, August 1969.

Hollanda, P. A., Scott, F., and Harabedian, A., "The Informative Value of Sampled Images as a Function

of the Number of Scans per Scene Object and the Signal to Noise Ratio," Photographic Science and Engineering, Vol. 14, No. 6, November-December 1970.

Johnston, Dorothy M., "Target Recognition on TV as a Function of Horizontal Resolution and Shades of Gray," Human Factors, Vol. 10 (3), pp. 201-210, 1968.

Legault, R. P., "Photoelectronic Imaging Devices,"(Biberman and Nudelman), Chapter 4, Vol. I, Plenum Press, New York, 1971.

Linfoot, E. H. "Transmission Factors and Optical Design," J. Opt. Soc. Am., Vol. 46, No. 9, September 1956.

Linfoot, E. H., "Quality Evaluations of Optical Systems," The University of Cambridge, England, Optical Acta, Vol. 5, Nos. 1-2 (March-June 1958).

Linfoot, E. H., "Optical Image Evaluation," University of Cambridge, England, Focal Press, London and New York (1966).

Mertz, P. and Gray, T., "The Theory of Scanning and its Relation to the Characteristics of the Transmitting Signal in Telephotography and Television," Bell System Tech. J. 13, p. 464, 1934.

Rose, A., "The Sensitivity Performance of the Human Eye on an Absolute Scale," J. Opt. Soc. Am. Vol. 38, No. 2, pp. 196-208, February 1948.

Rosell, F. A., "Noise Limitations to the Detection of Isolated Square Images on a TV Monitor," Westinghouse Aerospace, Baltimore, October 1968.

Scott, F., "Three-Bar Target Modulation Detectability," Photographic Science and Engineering, Vol. 10, No. 1, pp. 49-52, 1966.

Scott, F., "A Line-Scan Image Generator," Photographic Sci. and Eng., Vol. II, No. 5, pp. 348-351, September 1967.

Scott, F., "The Search for a Summary Measure of Image Quality--A Progress Report," Photographic Sci. and Eng., Vol. 12, No. 3, May-June 1968.

Schade, O. H. Sr., "A Method of Measuring the Optical Sine Wave Spatial Spectrum of Television Image Display Devices," J. SMPTE, 67, September 1958, p. 561-566.

Schade, O. H. Sr., Patent 3,030,440, Vertical Aperture Correction.

Schade, O. H. Sr., "Electro-Optical Characteristics of Television Systems," RCA Review, p. 5-37, March 1948. Part I - Characteristics of Vision and Visual Systems, March 1948. Part II- Electrooptical Specifications for Television Systems, June 1948. Part III - Electrooptical Characteristics of Camera Systems, September 1948, Part IV - Correlation of Imaging Systems.

Schade, O. H. Sr., "Image Gradation, Graininess and Sharpness in Television and Motion Picture Systems,"J. SMPTE. Part I - Image Structure and Transfer Characteristics, p. 137-171, Feb. 1951. Part II- The Grain Structure of Motion Picture Images--An Analysis of Deviations and Fluctuations of the Sample Number, Vol. 58, p. 181-222, March 1952. Part III - The Grain Structure of Television Images, Vol. 61, p. 97-164, August 1953.

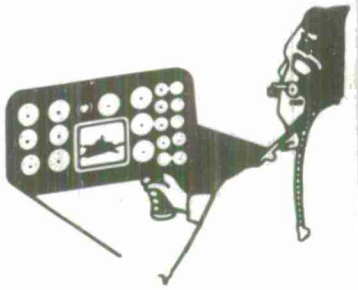
Schade, O. H., Sr., "Optical and Photoelectric Analog of the Eye," J. Opt. Soc. Am., Vol. 46, No. 9, p. 721-739, September 1956.

Schade, O. H. Sr., "An Evaluation of Photographic Image Quality and Resolving Power," J. SMPTE, Vol. 73, No. 2, p. 81-119, February 1964.

Schade, O. H., Sr., "The Resolving Power Functions and Quantum Processes of Television Cameras," RCA Review, September 1967, Vol. 28, No. 3.

Steedman, W. C. and Baker, C. A., "Target Size and Visual Recognition," Human Factors, August 1960.

Thompson, F. T., "Television Line Structure Suppression," Scientific Paper 8-1041-p14, Westinghouse Research Labs., Pittsburgh, Pa., 27 June 1957.



# ANALYSIS OF IMAGE-DETECTING AND DISPLAY SYSTEMS

Alvin D. Schnitzler
Institute for Defense Analyses
Arlington, Virginia

OFFICE OF NAVAL RESEARCH
TARGET ACQUISITION SYMPOSIUM
NAVAL TRAINING CENTER, ORLANDO, FLORIDA / 14,15,16 NOVEMBER 1972

# ANALYSIS OF IMAGE-DETECTING AND DISPLAY SYSTEMS

Alvin D. Schnitzler
Institute for Defense Analyses
Arlington, Virginia

## ABSTRACT

The operational performance of an image-detecting and display system is measured by the probability of acquiring a target and the false alarm probability. Performance depends on a large number of factors concerned with the scene, the atmosphere, the electro-optical imaging system, the visual system, and the search process automatically followed by the user. Motivated by knowledge of the search process and the structure and operation of the visual system the probability of acquiring a target can be decomposed into two factors, viz., the foveal imaging probability and the foveal detection probability. Both are subject to analysis by extension of the standard techniques of statistical communication theory. Although this analytical method is very general, as an example, it is applied to a low light level television system.

## INTRODUCTION

Correctly predicting the operational performance of image-detecting systems and choosing the most direct and efficient means to improve on it, depends on the development of a rigorous analytical expression relating operational performance to target and background characteristics and to system parameters.

Image-detecting systems generally fall into three categories:

- The unaided visual system.
- The visual system optically aided by a binocular or a telescope.
- The visual system electro-optically aided by an image intensifier, a television or an infrared converter system.

The principal components of the electrooptically aided visual system are indicated in Fig. 1. First, we have the objective for collecting radiant power and forming an image on the primary sensor and which with its primary lens or mirror generally serves as the aperture stop, determining the collection efficiency by the solid angle it subtends at the target. Next, we have the primary sensor which functions as a radiant power-to-electrical excitation transducer and which serves as the field stop of the system, determining the angular field-of-view by the solid angle which its image in object space subtends at the objective. Then we have the display for converting electrical power into luminous power and which in conjunction with the viewing distance determines the utilized angular field-of-view of the visual system by the solid angle which it subtends at the entrance pupil of the eye. The subjective magnification is the ratio of the utilized linear angular field-of-view of the visual system in display space to the linear angular field-of-view of

the electrooptical imaging system in object space. Next, we have the visual system for collecting luminous power from the display and forming an image on the retina where electrochemical processes are initiated in the receptor cells by the incident luminous power. The collection efficiency of the visual system is determined by the solid angle subtended at the display by the entrance pupil of the visual system. If an ocular is employed as in an image intensifier system, for example, it functions as a simple magnifier, increasing the collection efficiency and the utilized angular field-of-view of the visual system in display space. Finally, we have within the retina neurons which function as threshold devices transmitting electrical signals on to a higher level for processing if the electrical excitations received from their photoreceptor fields exceed their decision criteria.

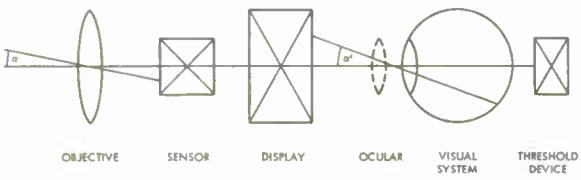



Figure 1. Diagram of Composite Electro-Optical-Visual System for Image Detection

The purpose of an electrooptical imaging system is to increase the acquisition and flow of information from a scene to the brain of a user over what would occur if he were forced to rely on his visual system alone. When the purpose is stated in these terms, an electrooptical imaging system is recognized as a communication system. Hence, the measures of performance, the methods of analysis and the methods of testing of electrooptical imaging systems are suggested by consideration of electrical communication systems.

A powerful tool which has been developed to analyze the performance of electrical communication system is provided by statistical communication theory. The general objective of statistical communication theory is to develop a statistical description of messages and noise in order to predict performance measured by the detection and false alarm probabilities. A more limited objective is often pursued in which it is assumed that a known message is being transmitted, but reception is degraded by the presence of noise. Hence, in this case, a statistical description of only the noise is required in order to predict the detection and false alarm probabilities.

The definition of the performance measures of an image-detecting system depend on the visual task described, but in general they are defined by the target acquisition and false alarm probabilities. The target acquisition task may range from merely detecting the presence of a target in an unstructured scene to precise identification of a particular target among a set of closely similar targets in a cluttered scene. In any case target acquisition is accomplished with an electrooptical imaging system by visually examining the display upon which the angular size and the contrast of the image of the target are larger than that of the target viewed directly by the unaided visual system.

Clearly, to develop a rigorous analytical expression relating operational performance to target and background characteristics and to system parameters a large number of factors must be considered:

- The nature of the radiant power transmitted from the scene - especially its spectral content.
- The spatial variation in reflection and/or emission of radiant power by targets and backgrounds to form the optical signal transmitted from the scene.

- The attenuation of the optical signal by absorption and scattering of radiant power in the intervening atmosphere.
- The attenuation of the optical signal due to the limited aperture size and aberrations of the objective of the image-detecting system.
- The efficiency and fidelity of the conversion of the optical signal into an electrical signal and the generation of electrical noise in the electrooptical imaging system.
- The efficiency and fidelity of the conversion of the electrical signal into a luminous signal at the display of the electro-optical imaging system.
- The attenuation of the optical signal and noise due to the limited aperture size and aberrations of the visual system.
- The technique employed by the user to acquire the images of targets of interest reproduced on the display which in turn depends on the complexity of the scene and the anticipated size and contrast of the target.

In order to simplify the problem of developing a rigorous analytical expression relating operational performance to all of the above factors, it will be convenient to decompose the probability of acquiring a target into two independent factors. Such a decomposition is motivated by the knowledge that the probability of detecting a target of size and contrast near the detection threshold at ordinary values of display luminance is much greater if the image of the target falls on the foveal region of the retina. In fact, foveal vision is so much more sensitive than extra-foveal vision that the muscles controlling the eye involuntarily rotate the eyeball from one fixation point to another in nearly discrete steps allowing approximately three fixations per second. Thus, the process of acquiring a target near the sensitivity threshold by the visual system consists of a series of fixations with a relatively small angular aperture corresponding to the fovea at a series of points within the larger solid angle corresponding to the display at the rate of approximately three fixations per second. According to this process the probability of acquiring a target depends on the probability that in a series of fixations on points of the display the target image falls on the fovea and on the probability that whenever the target image falls on the fovea, it will be detected. The process is described by

$$P_A(t) = P_S(t)P_D \qquad (1)$$

where $P_A(t)$ and $P_S(t)$ are the acquisition probability and the probability that the target image falls on the fovea, respectively, in a series of fixations of total duration $t$ and $P_D$ is the probability that the target image on the fovea is detected. As the size and contrast of the target increase the probability of extra-foveal detection becomes significant so that only one fixation is required to acquire the target and $P_S(t)$ as well as $P_D$ approaches unity.

It may be noted that Eq. (1) differs with Bailey's Eq. (1) describing his target acquisition model (Bailey, 1970) by omitting a conditional probability factor attributed to recognition once detection has occurred and an overall degradation factor attributed to noise. The first factor is omitted because recognition results from detection of details of target shape or structure which serve to establish

recognition of the target. Likewise, identification results from detection of smaller details of target shape or structure which serve to establish identification of the target. Obviously, the set of target details required for recognition or identification will depend on the target set to which a particular target belongs. If the target details required for recognition or identification are of sufficient size and contrast to permit a reasonable detection probability, detection of the target as a whole does not occur. Rather, at the primary decision level in the retina detection of the set of target details occurs and a corresponding set of signals is transmitted to the higher decision levels in the brain where the signal processing and comparison with memory required for recognition or identification occur. Mere detection of the target per se is irrelevant in this process. Thus, recognition or identification are not conditional on detection of the target and if the visual task is recognition or identification, it should be realized that in Eq. (1) $P_D$ symbolizes the detection of the requisite details. However, to avoid confusion either the symbol $P_D$, $P_R$ or $P_I$ can be used in Eq. (1) to designate the probability either of detecting the entire target, of detecting details characteristic of recognition, or of detecting details characteristic of identification, respectively.

Omission of the overall degradation factor attributed to noise by Bailey is made because the effect of noise is properly accounted for in the detection probability. Indeed, the detection probability is determined by the signal-to-noise ratio in conjunction with the decision criterion at the primary decision level in the neurons of the retina. In the absence of noise, no matter how small in size and contrast a target were found, the detection probability would be unity.

Now that the acquisition probability has been logically decomposed into two factors, each factor in turn will be analyzed in detail.

## THEORY

### Foveal Imaging Probability

The derivation of the cumulative probability that in a series of fixations of total duration $t$, the image of a target falls on the foveal region of the retina is essentially contained in Davies' derivation of the equations of glimpse theory (Davies, 1968). However, to clarify concepts relevant to the above description of the target acquisition process, a brief derivation of the foveal imaging probability $P_S(t)$ is presented here.

In order to derive an expression for the foveal imaging probability in accordance with the above description of the target acquisition process it is convenient to let $p_k$ be the probability that the target image falls on the fovea in the $k$th fixation. Then $(1 - p_k)$ is the probability of not imaging the target on the fovea in the $k$th fixation.

If the size and contrast of the target are relatively small so that the probability of detection when the image falls outside the fovea is nil, then it is reasonable to assume that the series of fixations points will be randomly distributed over the display. Moreover, experimental results tend to be in conformity with random fixation points (Davies, 1968). In this case the cumulative probability of not imaging the target on the fovea in $n$ fixations is given by the product $\prod_{k=1}^{n} (1 - p_k)$. Thus the cumulative foveal imaging probability in $n$ fixations is given by

$$P_S(n) = 1 - \prod_{k=1}^{n} (1 - p_k) \qquad (2)$$

It is reasonable to assume that the foveal imaging probability per fixation is a constant, i.e., that $p_k = p$ independently of the number of the fixation in the series. In addition if the duration of a fixation is designated by $T$ and the elapsed search time by $t$, then $n = t/T$ and the cumulative foveal imaging probability is given by

$$P_S(t) = 1 - (1 - p)^{t/T} \qquad (3)$$

The quantity $(1 - p)^{t/T}$ can always be expressed in exponential form by setting it equal to $\exp(u)$ and solving for $u$ which is found to be given by $u = (t/T)\ln(1 - p)$. Thus, Eq. (3) can be written as

$$P_S(t) = 1 - \exp(-t/\tau) \qquad (4)$$

where $\tau = -T/\ln(1 - p)$.

If the solid angle subtended by the image of the fovea on the display and the solid angle subtended by the display at the entrance pupil of the eye are designated by $\omega$ and $\Omega$, respectively, then $p = \omega/\Omega$. In addition if $\omega$ is small compared to $\Omega$, then $\ln(1 - p) \simeq - p$ and $\tau \simeq \Omega T/\omega$. Finally, in conformity with this approximation the cumulative foveal image probability is given by

$$P_S(t) = 1 - \exp(-\omega t/\Omega T) \qquad (5)$$

As an example, suppose the display subtends approximately 15 degrees or $6 \times 10^{-2}$ steradians. Now it is known that the most sensitive part of the fovea the fovea centralis subtends approximately $0.9^{\circ}$ and the entire fovea approximately $7^{\circ}$. Thus, the portion of the fovea employed in the target acquisition process will range from $0.9^{\circ}$ to $7^{\circ}$ depending on the size and contrast of the target. For this example let us assume that the fovea subtends $3^{\circ}$ or approximately $2 \times 10^{-3}$ steradians and the allowed search time is 10 seconds. Since, in addition, the duration of a fixation is well-known to be approximately 1/3 seconds, $P_S = 0.632$.

Rather than choosing the search time and solving for the foveal imaging probability it is interesting to choose the foveal imaging probability and solve for the search time. The expression for the search time is given by

$$t(P_S) = (-\Omega T/\omega)\ln(1 - P_S) \qquad (6)$$

In the above example if the required foveal imaging probability is chosen to be 0.90 or 0.99, the required search time is found to be 23 sec. or 46 sec., respectively. However, before drawing any conclusions concerning optimum angular size of the display from this example it is necessary to analyze the foveal detection probability.

## Foveal Detection Probability

In order to derive the foveal detection probability it is necessary to define the meaning of a statistical unit in an image-detecting system. A statistical unit is an aggregate of quanta or electrons which act together as a single unit or packet. In an electrooptical imaging system, at the primary sensor the statistical units are the primary electrons excited by the incident radiant power collected from the scene. Each primary electron is excited at a point in space and time (of course, only to the precision permitted by the Heisenberg uncertainty principle). At the display the statistical units are the aggregates of quanta whose mean number is equal to the mean value of the particle gain or intensification gain between the primary sensor and the display. In transmission from the primary sensor to the display and within the display the statistical units undergo both spatial and temporal dispersion so that they now occupy a small volume in space. In transmission from the display to the retina and within the retina the statistical units undergo further dispersion in space and time. At the retina the statistical units are the aggregates of primary molecular excitations whose mean value is the product of the mean value of a statistical unit emitted at the display, the collection efficiency of the visual system and the responsive quantum efficiency of the visual system.

If the magnitude of a statistical unit at the retina is greater than unity, an image-detecting system is said to be background limited due to shot noise or generation-recombination noise introduced at the primary sensor as a function of fluctuations in the background radiance of the scene. This condition prevails in the case of the unaided visual system at all luminance values except near zero where internal noise seems to be dominant and in the cases of the low light level television, the image

intensifier and the infrared converter systems at target and background conditions where they are mostly used. Under some conditions image-detecting systems such as television systems incorporating video amplifiers may not be background limited even though the magnitude of a statistical unit at the retina is greater than unity. This situation exists where no or insufficient intensifier gain is provided between the primary sensor and the video preamplifier. Such systems are said to be video preamplifier noise limited.

The magnitude of a statistical unit cannot be less than unity. Hence, if the product of the magnitude of a statistical unit at the display, the collection efficiency of the visual system and the responsive quantum efficiency of the visual system is less than unity, the product represents the fraction of the primary electrons which are detected as statistical units at the retina. In this case the image-detecting system is said to be display limited due to noise generated in the retina as a function of fluctuations in the background luminance of the display.

While the structure of an image at the retina is composed of statistical units, the individual statistical units are rarely perceived except when a very low radiant power scene is viewed through an image intensifier system, a low light level television system or when the totally dark adapted eye is suddenly exposed to a room light filtered through closed eyelids. Under most conditions, the rate of arrival of statistical units at the retina per unit area per second and their overlap in space and time due to spatial and temporal dispersion in both the electro-optical and visual systems is too great for individual statistical units to be perceptible. In addition electrical noise added particularly in video amplifier systems will partially or totally obscure the perception of individual statistical units. Although individual statistical units are rarely perceptible, they are the basic units which compose an image.

To derive the foveal detection probability it is assumed that the target or a detail of the target is a patch of uniform differential radiance relative to its local surroundings superimposed on a background radiance nominally equal to the average radiance of the scene. Furthermore, it is assumed that the boundary of the patch can be approximated by a simple geometrical figure such as a circle or a rectangle. If the objective optics were perfect, the temporal mean value of the irradiance in the perfect image formed on the primary sensor would be given by

$$\bar{n}_I(x,y) = \bar{n}_{BI} + \Delta\bar{n}_I(x,y) \qquad (7)$$

where $\bar{n}_{BI}$ is the temporal mean value of the background irradiance and $\Delta\bar{n}_I(x,y)$ is the temporal mean value of the differential irradiance in the perfect image relative to the background. Since it is assumed that the radiance of the patch is uniform, $\Delta\bar{n}_I(x,y)$ is a constant within the boundary of the perfect image and zero elsewhere.

The temporal mean value of the number of electrons $\bar{N}_I$ excited at the primary sensor by the incident radiant power within the area $a_I$ of the perfect image and during an exposure or integration time $\tau_s$ is obtained by multiplying Eq. (7) by $\eta_s a_I \tau_s$ where $\eta_s$ is the responsive quantum efficiency of the primary sensor. However, in a series of exposures occurring at $t_1$, $t_2$, ..., $t_k$, ... the number of excited electrons $N_I(t_k)$ is a random variable with a Poisson probability density. Thus, even though the target were absent, the number of electrons excited within the area of the perfect image would deviate from the temporal mean value and would be given by

$$N_I(t_k) = \bar{N}_{BI} + \Delta N_I(t_k) \qquad (8)$$

where $\Delta N_I(t_k)$ is the value of the random deviation in the kth exposure.

If the temporal mean value is not too small, the Poisson probability density of a random variable is indistinguishable from the more general

Gaussian probability density - the variance in this special case being equal to the mean value. Thus, in the absence of a target the probability density of the random deviation in the number of electrons excited per area $a_I$ and exposure $\tau_s$ is given by

$$p(\Delta N_I)d(\Delta N_I) = (2\pi)^{-1/2}\sigma_B^{-1} \times \exp[-(\Delta N_I)^2/2\sigma_B^2]d(\Delta N_I) \quad (9)$$

where $\sigma_B$ is the standard deviation equal to $(\bar{N}_{BI})^{1/2}$.

From Eq. (9) the normalized Gaussian probability density is obtained by letting $x = \Delta N_I/\sigma_B$ to obtain

$$p(x)dx = (2\pi)^{-1/2}\exp(-x^2/2)dx \quad (10)$$

If the target is present the deviation in the number of excited electrons in an exposure is approximately equal to the algebraic sum of the random deviation due to the background and the differential number due to the temporal mean value of the differential irradiance of the perfect image relative to the background. This statement is based on the assumption that the differential irradiance in the vicinity of the threshold for detection is small compared to the background. Later it will be shown that this is a valid approximation for real images at the decision level in the output of the visual system. By the above statement if the target is present, the random deviation in the number of excited electrons in the kth exposure is given by $\Delta N_I(t_k) - \Delta\bar{N}_I$ and its probability density is given by

$$p(\Delta N_I - \Delta\bar{N}_I)d(\Delta N_I) = (2\pi)^{-1/2}\sigma_B^{-1} \times \exp[-(\Delta N_I - \Delta\bar{N}_I)^2/2\sigma_B^2]d(\Delta N_I) \quad (11)$$

This equation may be expressed as a normal Gaussian probability density by substituting $k_I = \Delta\bar{N}_I/\sigma_B$ and again letting $x = \Delta N_I/\sigma_B$ to obtain

$$p(x-K_I)dx = (2\pi)^{-1/2} \times \exp[-(x-k_I)^2/2]dx \quad (12)$$

where $\Delta\bar{N}_I$ and $k_I$ are recognized as the signal and the signal-to-noise ratio, respectively, of the electron excitations at the primary sensor due to the target.

If an image-detecting system were perfect introducing no additive electrical noise and no temporal or spatial dispersion and forming a perfect image composed of statistical units which were points in space and time at the retina, then the output signal-to-noise ratio $k_o$ at the decision level would equal the signal-to-noise $k_I$ at the primary sensor. In this case the foveal detection and false alarm probabilities could be derived from the normal probability density equations (10) and (12).

Although in practice spatial and temporal dispersion occurs in each component of the image-detecting system including the visual system so that at the retina the real image area is greater than the perfect image area and the statistical units have a finite area and duration, the probability density of the random deviations at the output of the visual system remains Gaussian. Hence, the foveal detection and false alarm probabilities can be derived from Eqs. (10) and (12) if $k_I$ is replaced by $k_o$ and an expression is derived for $k_o$ in terms of target and background characteristics and system parameters.

At the decision level in the output of the visual system a neuron fires sending a message over one of the fibers of the optic nerve if and only if the signal and/or noise levels exceed a threshold level known as the decision criterion. If $\Delta N_o(t_k)$ is the differential number of molecular excitations produced in the retina by the algebraic sum of the random deviation in the number due to the background in the kth exposure and the temporal mean value of the number due to the difference in radiance between the target and background, then a neuron fires if and

only if $\Delta N_o(t_k) \geq \Delta N_T$ where $\Delta N_T$ is the decision criterion. If we divide by the standard deviation $\sigma_o$ this conditional equation becomes $x \geq k_T$ where $k_T = \Delta N_T/\sigma_o$ is the decision criterion expressed as a threshold signal-to-noise ratio.

The relations between the normalized Gaussian probability densities and the decision criterion are illustrated in Fig. 2 where $k_T$ and $k_o$ have been chosen equal to 2 and 3, respectively. The false alarm probability is equal to the area in Fig. 2 under the $p(x)$ curve bounded on the left by $x = k_T$. Thus, we have

$$1 - P_F = \int_{-\infty}^{k_T} p(x)dx \qquad (13)$$

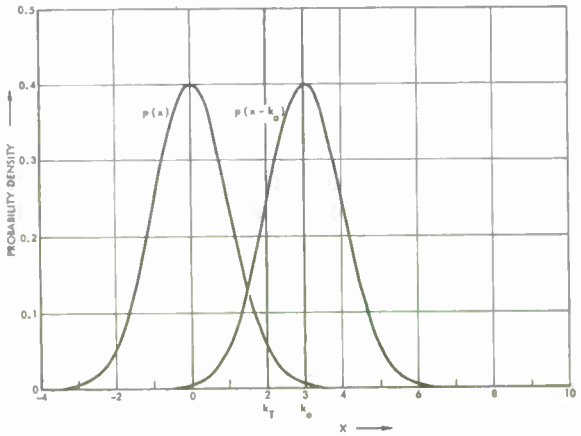



Figure 2. Noise and Signal plus Noise Normal Gaussian Probability Density Functions; $k_T$ is Threshold Signal-to-Noise Ratio; $k_o$ is Output Signal-to-Noise Ratio

This integral is recognized as the standard tabulated error function. Hence,

$$1 - P_F = \text{erf}(k_T) \qquad (14)$$

The detection probability is given by the area in Fig. 2 under the $p(x-k_o)$ curve bounded on the left by $x = k_T$, i.e.,

$$P_D = \int_{k_T}^{\infty} p(x-k_o)dx \qquad (15)$$

Since $p(x-k_o)$ is symmetrical about the $x = k_o$ axis, this area is the same as the area under the $p(x-k_o)$ curve bounded on the right by $k_o + (k_o - k_T)$. Hence,

$$P_D = \int_{-\infty}^{2k_o-k_T} p(x-k_o)dx \qquad (16)$$

To express $P_D$ as an error function we let $u = x - k_o$. Then we have

$$P_D = \int_{-\infty}^{k_o-k_T} p(u)du \qquad (17)$$

or

$$P_D = \text{erf}(k_o-k_T) \qquad (18)$$

It can be shown (Schnitzler, 1973) by comparison of Eq. (18) with experimental statistical detection data reported by Blackwell (1963) that the value of the threshold signal-to-noise ratio falls in the interval (2, 3) for a range of background luminance from $10^{-6}$ to $10^2$ ft.-L and a range of target angular sizes from essentially a point to $10^2$ mr. The facts that the value of $k_T$ deduced from experimental data is reasonable and essentially invariant to changes in background luminance and target size strongly substantiates the above description of the decision process.

To complete the derivation of the foveal detection probability it remains to derive an expression for the output signal-to-noise ratio $k_o$ in terms of target and background characteristics and system parameters. The effects of spatial and temporal dispersion in the components of an image-detecting system are to increase the area of the target image and to increase the area and duration of a statistical unit.

Although the area of the target image is increased by spatial dispersion, the differential number of statistical units within the image area remains unchanged. Since the differential number of statistical units is unaffected by spatial dispersion, the output signal is equal to the product of the magnitude of a statistical unit at the retina and the number of statistical units due to the differential irradiance in the perfect image of the target at the primary sensor. The magnitude of a statistical unit at the retina is $(G_D \zeta \eta_E)$ where $G_D$ is the particle gain between the primary sensor and the display, $\zeta$ is the collection efficiency and $\eta_E$ is the responsive quantum efficiency of the visual system. The differential number of statistical units in the perfect image at the primary sensor is equal to $(\eta_S \Delta\bar{n}_I a_I \tau_S)$, where $\eta_S$ is the responsive quantum of the primary sensor, $\Delta\bar{n}_I$ is the mean differential irradiance as defined in Eq. (7), $a_I$ is the area of the ideal image of the target on the primary sensor and $\tau_S$ is the effective integration time of the image-detecting system. Thus, the output signal $S_o$ is given by

$$S_o = (G_D \zeta \eta_E) \eta_S C_I \bar{n}_{BI} a_I \tau_S \qquad (19)$$

where $C_I \bar{n}_{BI}$ has been substituted for $\Delta\bar{n}_I$ and $C_I$ is the contrast of the ideal image or of the target itself.

While the differential number of statistical units contained in the image of a target is unaffected by spatial dispersion, the number of background statistical units increases in proportion to the area of the image. At the same time, the increase in area and duration of statistical units by spatial and temporal dispersion results in an overlapping of statistical units. The increase in number of background statistical units tends to increase the standard deviation while the overlapping of statistical units tends to decrease the standard deviation in the number of molecular excitations produced within the area of the target image per integration time.

In order to properly account for the effects of spatial and temporal dispersion on noise in an image-detecting system, it is necessary to represent the random excitation of electrons at the primary sensor and the response at the output of the visual system by power density spectra. It is shown in text books concerned with statistical communication theory (Lee, 1961) that the noise power density spectrum $\Phi_i$ due to a Poisson distribution of unit impulses which represent the random excitation of electrons at the primary sensor is given by

$$\Phi_i = \eta_S \bar{n}_{BI} \qquad (20)$$

Likewise, if the image-detecting system is linear and both spatially and temporally invariant, the output noise power density spectrum is equal to the product of the square of the magnitude of the transfer function and the noise power density spectrum at the input. Thus, the output noise power density spectrum is given by

$$\Phi_o(\nu) = (G_D \zeta \eta_E)^2 T_D^2 T_E^2 R_E^2 \Phi_i(\nu) \qquad (21)$$

where the frequency $\nu$ consists of three components, two for the two orthogonal axis in the plane and one for the temporal frequency, $(G_D \zeta \eta_E)$ is the magnitude of a statistical unit defined above, $T_D$ is the modulation transfer function of the system from the primary sensor to the display, $T_E$ is the modulation transfer function of the visual system and $R_E$ is a low pass filter function which accounts for spatial integration over the target image by the visual system.

The variance of the response rate averaged over the image area is obtained by integrating the output power density spectrum over all frequencies. Thus, the variance is given by

$$\sigma_o^2 = (G_D \zeta \eta_E)^2 \eta_S \bar{n}_{BI} \int_{-\infty}^{\infty} T_D^2 T_E^2 R_E^2 d\nu \qquad (22)$$

Since the system is invariant in time the temporal frequency component can be separated from Eq. (22) and performed separately. The result is

$\int_{-\infty}^{\infty} T_D^2(f) T_E^2(f) df = 1/\tau_S$ where the integral is the noise equivalent temporal frequency bandwidth and $\tau_S$ is the response time of the system.

If the area $a_o$ of the target image at the retina is large compared to the area of a statistical unit, the total variance of the random excitations is equal to the product of the variance of the average rate of response, the square of the image area and the square of the response time. This result is given by

$$\sigma_o^2 = (G_D^2 \zeta \eta_E)^2 \eta_S \bar{n}_{BI} a_o \int_{-\infty}^{\infty} T_D^2 T_E^2 R_E^2 d\nu \qquad (23)$$

If Eq.'s (19) and (23) are combined, the output signal-to-noise ratio is found to be

$$k_o = C_I (a_I/a_o)[\eta_S \bar{n}_{BI} \tau_S / \int_{-\infty}^{\infty} T_D^2 T_E^2 R_E^2 d\nu]^{\frac{1}{2}} \qquad (24)$$

Since the output signal-to-noise ratio is not accessible for direct measurement in the visual system, it is conventional to report experimental data in the form of input contrast required for 50 percent detection probability as a function of background radiance and target angular size. Since the input contrast required for 50 percent detection probability corresponds to equality between the output and threshold signal-to-noise ratios, the expression noise required input contrast is adopted to designate it. The noise required input contrast $C_{IN}$ is obtained by setting $k_o = k_T$ in Eq. (24) and solving for $C_I$. The result is given by

$$C_{IN} = m_S(\alpha_o^2/\alpha^2)\left[\int_{-\infty}^{\infty} T_D^2 T_E^2 R_E^2 d\nu\right]^{\frac{1}{2}} \qquad (25)$$

where

$$m_S = 4k_T / D_S(\pi \eta_S \bar{n}_B \tau_S)^{\frac{1}{2}} \qquad (26)$$

In going from Eq. (24) to (25) the shape of the target was assumed to be a circular disc in anticipation of comparison with experimental data (Blackwell, 1946), the circular areas $a_o$ and $a_I$ have been expressed in terms of their angular subtenses $\alpha_o$ and $\alpha$, respectively, at the entrance pupil of the system, the background irradiance $\bar{n}_{BI}$ at the primary sensor has been expressed in terms of the background exitance $\bar{n}_B$ at the scene and in Eq. (26) $D_S$ is the diameter of the entrance pupil of the system.

Since the results of a reliable determination of the modulation transfer function of the visual system are not available, it is useful to express the noise required input contrast of an image-detecting system in terms of the noise required input contrast of the unaided visual system viewing a projector screen at the same luminance as the display of the electrooptical system. The result is given by

$$C_{IN}(\bar{E}) = C_{IN}(\bar{E}_D, M)\, M(m_S/m_E) \times \left[M^2\alpha^2 + \alpha_S^2)/(M^2\alpha^2 + \alpha_E^2)\right] \times \left[\int_{-\infty}^{\infty} T_D^2 T_E^2 R_E^2 d\nu \Big/ \int_{-\infty}^{\infty} T_R^2 R_E^2 d\nu\right]^{\frac{1}{2}} \qquad (27)$$

where $\bar{E}$ is the mean luminance of the scene, M is the magnifying power of the electrooptical system, $m_E$ is obtained from Eq. (26) by replacing $D_S, \eta_S, \tau_S$ by the corresponding visual parameters and $\bar{n}_B$ by the mean exitance

$\bar{n}_D$ of the display, $\alpha_S$ is the angle subtended at the entrance pupil of the image-detecting system by the image of a point source on the retina scaled to its size in object space, $\alpha_E$ is the analogous parameter for the unaided visual system, $T_R$ is the retinal part of the modulation transfer function of the visual system and $C_{IN}(\bar{E}_D,M)$ is the noise required input modulation of the unaided visual system viewing a projector screen at the display luminance $\bar{E}_D$. This latter factor depends on M because the target angular size is with reference to the input to the image-detecting system.

To present a specific example a background limited low light level television system will be considered. To simplify the example it is assumed that modulation transfer function of the overall image-detecting system is limited by the retinal part of the visual system so that the integral terms in Eq. (27) cancel and the bracketed coefficient is unity. Thus, in this case Eq. (27) reduces to

$$C_{IN}(\bar{E}) = M(m_S/m_E)C_{IN}(\bar{E}_D,M) \qquad (28)$$

If reasonable values of the low light level television parameters are assumed such as the objective focal length equal to 12 in., the width of the primary sensor equal to 32 mm, the viewing distance 24 in., and the width of the display equal to 11.3 in., the magnifying power equals 4.5.

The separate effects of magnifying power, brightness gain, collection efficiency and responsive quantum efficiency on the performance of the low light level television system as an adjunct of the visual system are shown in Fig. 3. It is assumed that the targets are circular discs in order to make use of the Blackwell data for detection of circular discs projected on a screen at $10^2$ ft-L corresponding to the display luminance. In principle the procedure could be applied to any target if standard observer data for detection on a projector screen at the display luminance were available. The Blackwell experimental disc detection data is shown in Fig. 3 by the dashed curve labeled "Display Limit." This curve depicts the logarithm of the function $C_{IN}(\bar{E}_D,M)$ in Eq. (28). Note that the values of $1/\alpha$ are for object space of the low light level television system.

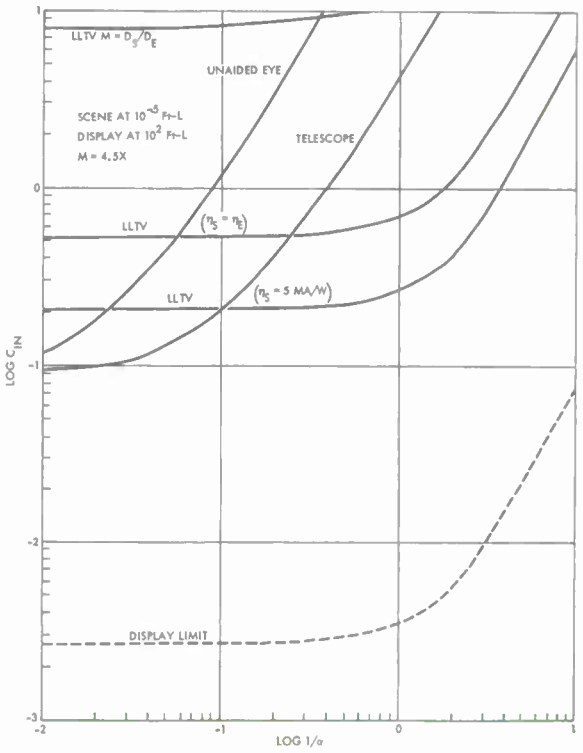



**Figure 3. Logarithm of Noise Required Input Contrast versus Logarithm of Reciprocal Target Angular Size where the Angle is in Milliradians. The Curves are Discussed in the Text**

If the unaided visual system were to view the scene which is at a luminance of $10^{-5}$ ft-L, then the Blackwell data results in the log $C_{IN}$ vs log $(1/\alpha)$ curve labeled "Unaided Eye" in Fig. 3. Now the effect of subjective magnification alone is shown by the curve labeled "Telescope" which is obtained from the unaided eye curve by increasing the values of $(1/\alpha)$ by the magnifying power 4.5.

If the diameter of the aperture of the low light television system were equal to that of the telescope but the brightness gain of $10^7$ were utilized, then the curve labeled LLTV, $M = D_S/D_E$ is obtained from the "Display Limit" curve by employing Eq. (28). Note that the performance of the unaided eye would be better for values of $1/\alpha$ less than approximately 0.4 $mr^{-1}$.

If the full aperture of the system were used the curve labeled ($\eta_S = \eta_E$) is obtained from the "Display Limit" curve by employing Eq. (28). Of course, enlarging the aperture improves the performance but the unaided eye still performs better at values of $1/\alpha$ less than approximately 0.06 $mr^{-1}$.

Finally, if in addition to subjective magnification and larger aperture the primary sensor of the low light level television system has a responsive quantum efficiency of 5 mA/W the curve labeled ($\eta_S$ = 5mA/W) is obtained. While the performance of the unaided eye is better at values of $1/\alpha$ less than 0.03 $mr^{-1}$ such large targets are of little interest in practice.

The effect of varying the size of the display and hence the subjective magnification of the assumed low light level television system is shown in Fig. 4 for a scene luminance of $10^{-3}$ ft-L. Note that as the subjective magnification decreases the performance at large angles improves but the performance at small angles degrades. At M = 0.59 the coefficient of $C_{IN}(\bar{E}_D, M)$ is unity. Further decreases in M result in degradation at all angles.

## DISCUSSION AND CONCLUSIONS

The explanation for the dependence of performance (contrast sensitivity) of the low light level television system on subjective magnification M at large target angular sizes as shown in Fig. 4 is that the area of the image of the target on the retina is greater than the maximum area over which the visual system can integrate. Rigorous analysis (Schnitzler, 1973) of the Blackwell disc detection data reveals that at a display luminance of 100 ft-L the maximum angular diameter $\alpha_1$ of the effective area of integration on the retina is 3.7 mr. Indeed, the reciprocal (0.28 $mr^{-1}$) corresponds approximately to the value of $1/\alpha$ at which the M = 1X curve in Fig. 4 bends upward. As M is made larger the angular size in object space at which the image of the target on the retina exceeds 3.7 mr is shifted to smaller values, i.e., the value of $1/\alpha_1$ with reference to object space increases in proportion to M. This is apparent in Fig. 4 by comparing the angles at which the upward bend in the $C_{IN}$ curves occurs.

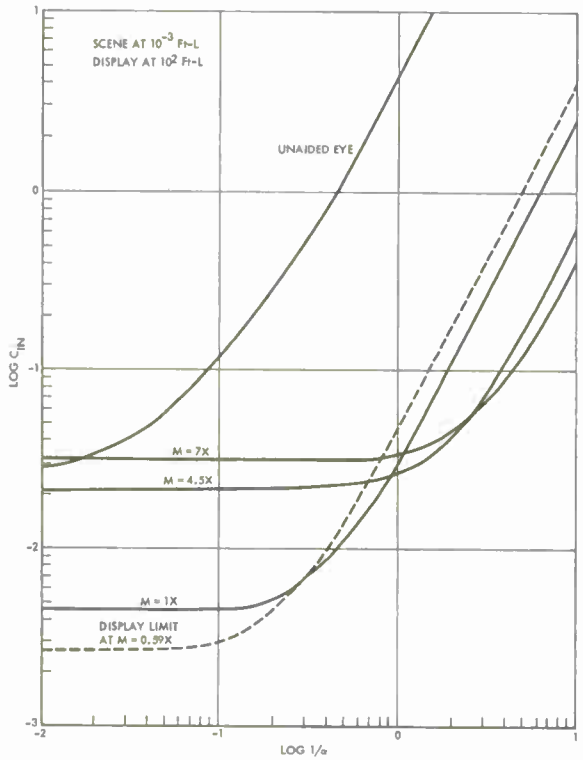



**Figure 4. Effect of Varying Subjective Magnification upon Log Noise Required Input Contrast versus Log Reciprocal Target Angular Size**

At large angular sizes such that the area of the image of the target on the retina exceeds the maximum integration area ($1/\alpha < 1/\alpha_1$ in object space) an increase in M results in a smaller fraction of the target image

filling the maximum integration area on the retina. Thus, the number of statistical units due to the differential radiance of the target relative to the background and the number of statistical units due to the background radiance which are imaged on the integration area both decrease in proportion to $M^2$. Since the signal and noise are proportional to the differential number of statistical units and the square root of the number of background statistical units, respectively, they are proportional to $1/M^2$ and $1/M$, respectively. Hence the S/N ratio is proportional to $1/M$ and the noise required input contrast at large values of target angular size is proportional to $M$.

The explanation for the dependence of performance on subjective magnification shown in Fig. 4 at small target angular sizes is that spatial dispersion determines a minimum image size on the retina. In this case it can be shown (Schnitzler, 1973) that the noise required input contrast, $C_{IN}(\bar{E}_D,M)$ in Eq. (28), of the unaided visual system viewing a screen at the display luminance $\bar{E}_D$ is proportional to $1/M^2\alpha^2$. Since it was shown by Eq. (28) that the noise required input contrast $C_{IN}(\bar{E})$ of the system is proportional to $MC_{IN}(\bar{E}_D,M)$, we have $C_{IN}(\bar{E}) \sim 1/M\alpha^2$. If a particular target of small angular size and some known contrast is considered, then $1/\alpha \sim M^{\frac{1}{2}}$. Since $1/\alpha$ is proportional to range, the detection range $R_D$ is proportional to $M^{\frac{1}{2}}$.

If the equations developed for the foveal imaging probability, the search time and the foveal detection probability are considered together, the dependence of performance on system parameters can be examined. One parameter which has received considerable attention recently is the size of the display.

It is of interest to consider the cost in search time of increasing the display size in order to achieve greater detection range. It was shown by Eq. (6) that for a given foveal imaging probability the required search time t is proportional to the solid angle $\Omega$ subtended at the eye by the display, i.e., $t \sim \Omega$. Since it was shown above that $R_D \sim M^{\frac{1}{2}}$ and if only the display size is variable, then $R_D \sim \Omega^{\frac{1}{4}}$. Thus, if the display size is variable, the search time is proportional to the fourth power of the detection range. For an 11.3 in wide display viewed from 24 in the search time would be 5 sec even if the entire 7 degree diameter of the fovea were effective. Hence, one should be exceedingly cautious about increasing the display size in order to obtain greater detection range unless random search can be avoided by prebriefing. If greater detection range is required, the necessary subjective magnification can be obtained by decreasing the field of view in object space.

To design a background limited electrooptical imaging system to be used in random search for low contrast targets, the first step is to choose as large a display as is consistent with allowable search time. Then make a trade-off between the desired detection range and the field of view in object space required to achieve the necessary subjective magnification for detection. If it is necessary to reduce the field of view in order to achieve sufficient subjective magnification for a required detection range, then it should be accomplished so as to avoid degradation of the modulation transfer function of the electrooptical system by choosing a longer focal length objective rather than by reducing the area of the primary sensor. If a required field of view and detection range cannot be simultaneously achieved, then a larger objective aperture must be chosen.

## REFERENCES


Bailey, H. H. (1970), Memorandum RM-6158/1-PR, The Rand Corporation.

Blackwell, H. R. (1946), J. Opt. Soc. Am. 36, 624; Summary Tech. Rpt. of Div. 16, Vol. 2, Natl. Def. Research Committee.

Blackwell, H. R. (1963), J. Opt. Soc. Am. 53, 129.

Davies, E. B. (1968), "Visual Search Theory with Particular Reference to Air-to-Ground Vision," Royal Aircraft Establishment, Tech. Rpt. 68055.

Lee, Yak Wing (1961), Statistical Theory of Communication, John Wiley and Sons, New York.

Schnitzler, Alvin D. (1973), "Determination of the Image Detection Parameters of the Visual System," to be published.

# SIGNAL-TO-NOISE RATIO THRESHOLDS FOR IMAGE DETECTION RECOGNITION AND IDENTIFICATION

by

Frederick A. Rosell and Robert H. Willson
Westinghouse Electric Corporation
Defense and Electronic Systems Center
Baltimore, Maryland 21203

OFFICE OF NAVAL RESEARCH
TARGET ACQUISITION SYMPOSIUM
NAVAL TRAINING CENTER, ORLANDO, FLORIDA / 14,15,16 NOVEMBER 1972

# SIGNAL-TO-NOISE RATIO THRESHOLDS FOR IMAGE DETECTION RECOGNITION AND IDENTIFICATION

Frederick A. Rosell and Robert H. Willson
Westinghouse Electric Corporation
Defense and Electronic Systems Center
Baltimore, Maryland 21203

Image signal-to-noise ratio models are hypothesized for the detection, recognition and identification of scene objects using an equivalent bar pattern approach. The extent to which the models apply will be discussed and the observer's threshold signal-to-noise ratio requirements are determined for various levels of object discrimination through psychophysical experimentation.

## INTRODUCTION

During the past few decades, a theory of imaging, based on the image signal-to-noise ratio concept, has been evolving through the efforts of (Barnes, 1932; DeVries, 1943; Rose, 1948; Shade, 1967; Rosell, 1969, etc.). Most of the effort has been devoted to determining an observer's ability to detect images of simple geometry and, rightfully so; simple images can be easily generated and can be mathematically described in both the space and the Fourier spatial frequency domain. By simple images, it is implied that the test images are of regular shape such as squares, rectangles or disks, or patterns which are periodic in one direction such as trains of sine or square waves and aperiodic in the other. In nearly all cases, the image backgrounds were uniform. During the past few years, the authors have attempted to further verify and to determine the limitations of the current theory and to apply the results to real imaging systems (Rosell and Willson, 1971, 1972). More recently, we have been attempting to extend the theory to more complex images such as may be encountered in aerial reconnaissance. A further goal is to correlate the detectability of the simple geometric images with more complex levels of image discrimination. This is of value since the simple images now used in the laboratory evaluation of sensors can be used as an indicator of the level of performance to be expected in field use.

In the case of simple imagery, it is sufficient that the observer "detect" the image or "resolve" a bar pattern structure. With more complex images, the observer may be required to perform a discrimination task such as differentiating between two classes of vehicles or identifying a particular type of aircraft. In this paper, we will briefly describe the elementary theory, review a small number of the experiments designed to verify the elementary theory and discuss in somewhat more detail the recent efforts to extend the theory to real imagery. Specifically, we will be concerned with levels of scene object discrimination ranging from the lowest level of detection to the more complex levels of recognition and identification.

## THE ELEMENTARY THEORY

The elementary theory regarding the detection of simple aperiodic and periodic images has been published previously (Rosell and Willson, 1970) and will be further developed and reviewed by R. Sendall in these Proceedings. For the present purpose, we note one of the important results which relates the signal-to-noise ratio, $SNR_{DI}$, of the image as it appears on the display of a television sensor to the video signal-to-noise ratio, $SNR_V$, as measured in the sensor's electrical

channel. This relation is

$$SNR_{DI} = [2t\Delta f_V(\frac{a}{A})]^{\frac{1}{2}} \bullet SNR_V \qquad (1)$$

where t is an integration time in sec., $\Delta f_V$ is the video bandwidth in Hz., and a is the image area relative to the total effective photosurface area on which a is imaged. This relationship holds only for images which are unlimited by the sensor's spatial frequency response, of uniform intensity amplitude and for noise spectrum which is essentially white in character.

The above equation describes the $SNR_{DI}$ obtainable from the sensor. If now, we could determine the $SNR_{DI}$ required by the observer to perform some simple task such as detecting a rectangle, we could determine some sensor performance limits by matching that obtainable with that required. To obtain the observer's requirements, we used the experimental set-up shown in Fig. 1. In this experiment, a

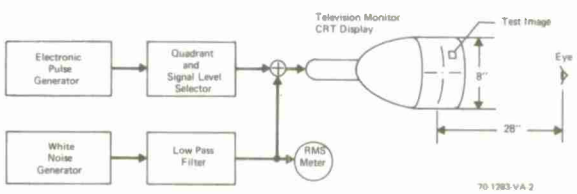



Fig. 1 The Display Signal-to-Noise Ratio Experiment.

signal pulse of rectangular waveshape but variable duration is electronically generated and mixed with band-limited white noise of Guassian distribution. The spatial image displayed on the cathode ray tube (CRT) display is a rectangle which can appear in any of four quadrants (but always in the same position in the quadrant selected). The observer is asked to specify the quadrant in which the image is located as the video signal-to-noise ratio and the image locations are randomly varied. The observer is asked to specify the image location whether he could see it or not. The probability of detection, determined in this manner, was then corrected for chance using the formula $P_d = (P_o - P_c)/(1 - P_c)$ where $P_d$ is the corrected probability, $P_o$ is the raw probability data and $P_c$ is the probability due to chance (0.25 in the case cited). Two noise bandwidths were used, 7.1 and 12.5 MHz and the observation times per trial were usually 10 seconds. The observer distance from the 8" high picture displayed was 28" unless otherwise specified and the display background luminance was either 0.2 - 0.3 or 1 ft-Lambert. The television monitor was operated at 30 frames per second with a 525 line scan in the vertical.

For the rectangle experiments, it was found to be convenient to define the image size in terms of the dimensions of a single scan line. Thus, we define the quantities $L_x$ and $L_y$ as

$$L_x L_y = (490)^2 \alpha(a/A) , \qquad (2)$$

where 490 is the number of active lines in a conventional 525 line television display and $\alpha$ is the width-to-height picture aspect ratio of the total effective picture on the CRT. Combining Eqs. (1) and (2), we have

$$SNR_{DI} = (1/490) \bullet (2L_x L_y t\Delta f_V/\alpha)^{\frac{1}{2}} \bullet SNR_V . \qquad (3)$$

This is the equation used to calculate the $SNR_{DI}$ for the rectangular images used in the experiments below. The numerical values used were t = 0.1 sec and $\alpha = 4/3$.

For the experiment, we hypothesized that the $SNR_{DI}$ required to liminally* detect a stationary rectangular image of variable length would be a constant independent of the image's area. To test this notion, we measured probability of detection vs $SNR_V$ and $SNR_{DI}$ for rectangles of size 4 x 4, 4 x 64, 4 x 128 and 4 x 180 scan lines. The results, plotted in the form of probability of detection vs video SNR are shown in Fig. 2 where it is seen that the larger the image length, the smaller the $SNR_V$ required for a given level of probability. When the probability of detection is plotted vs the display signal-to-noise ratio as calculated using

* For liminal detection, we mean a 50%, or threshold, probability of detection.

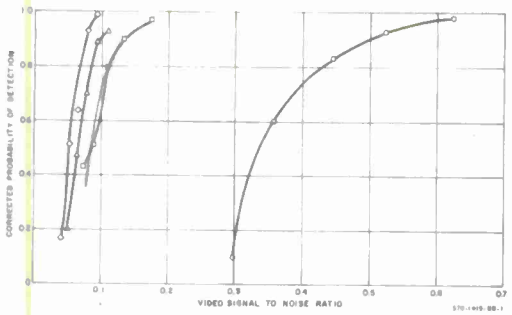



Fig. 2 Corrected Probability of Detection vs Video Signal-to-Noise Ratio for Rectangular Images of Size O 4 x 4, □ 4 x 64, Δ 4 x 128 and ◇ 4 x 180 Scan Lines.

Eq. (3) and shown in Fig. 3, the data collapses to a single curve confirming the original hypothesis. These results hold for a wide range of rectangles so long as not more than one of the rectangle's dimensions exceeds about 1/2° relative to the observer's eye. For a more detailed discussion of this experiment an and many others, the reader is referred to (Rosell, 1972). The above results represent 800 observations by 5 observers.

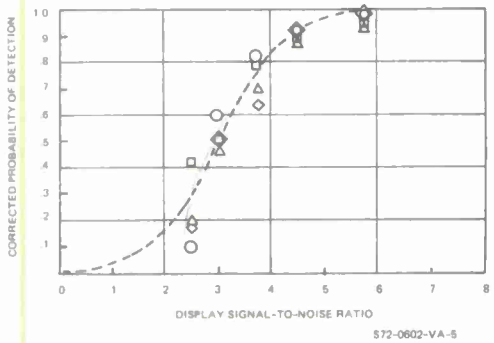



Fig. 3 Corrected Probability of Detection vs $SNR_{DI}$ Required for Rectangular Images of Size O 4 x 4, □ 4 x 64, Δ 4 x 128 and ◇ 4 x 180 Scan Lines.

Although the equations above were derived for an isolated rectangular image, it is hypothesized that they also apply to the detection of bar and sine wave patterns on the premise that to detect the presence of a bar pattern, the observer must detect the presence of a single bar. However, the threshold signal-to-noise ratio required to detect a bar in the presence of a number of bars may differ from that needed to detect an isolated bar on a uniform background but will not differ much as we will see. Suppose that a single bar in a pattern is of size $\Delta y/Y$ wide by $\epsilon\Delta y/Y$ long where $\epsilon$ is the bar length-to-width ratio and Y is the picture height. Note that we use dimensionless size units because of its convenience when a number of image size changes occur within a system. Also, for reasons which will become apparent as we progress, we will prefer to use the reciprocal dimension, N, as given by

$$N = \frac{Y}{\Delta y} \quad . \tag{4}$$

The units of N are "lines per picture height". Then, the bar image area relative to the total effective photocathode area is equal to

$$\left(\frac{a}{A}\right) = \frac{\epsilon \Delta y^2}{\alpha Y^2} = \frac{\epsilon}{\alpha N^2} \quad , \tag{5}$$

where $\alpha$ is the displayed horizontal to vertical picture aspect ratio and N is designated the pattern's spatial frequency in "lines per picture height". With this result, Eq. (3) becomes

$$SNR_{DI} = [2t\Delta f_V/\alpha]^{\frac{1}{2}} \frac{\epsilon^{\frac{1}{2}}}{N} \cdot SNR_V \quad . \tag{6}$$

To obtain the observer's requirements for the periodic patterns, we employed the experimental set-up of Fig. 4.

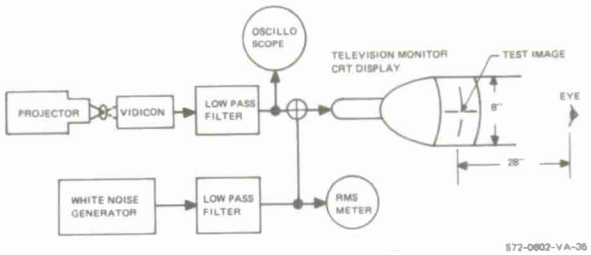



Fig. 4 Experimental Set-up for the Television Camera Generated Imagery.

The periodic test images used were primarily bar patterns, of various height-to-width ratios, spatial frequencies and numbers of bars in the pattern which were projected on the faceplate of a high resolution 1½" vidicon operated at high-light video signal-to-noise ratios of 50:1 or better. The camera and TV monitor were operated at 25 frames/second with 875 scanning lines (825 active). Band-limited white noise of Guassian distribution was mixed with the camera generated signal. Both the signals and noise were passed through identical filters (noise equivalent bandwidth of 12.5 MHz) prior to mixing in the monitor. The monitor luminance was approximately 1 ft. Lambert in most cases.

Because the vidicon camera has a limited modulation transfer function, the displayed images were square wave patterns at very low spatial frequencies, sine wave patterns at high spatial frequencies and distorted square wave patterns in between. Generally, the peak-to-peak signal is measured in the laboratory but as Schade (1967) notes, the displayed bar pattern detectability depends upon the mean signal rather than the peak-to-peak signal modulation. However, knowing the sensor's modulation transfer function $|R_o(N)|$, the mean signal response, called the square wave flux response $R_{SF}(N)$ can be computed from

$$R_{SF}(N) = \frac{8}{\pi^2} \sum_{k=1}^{\infty} \frac{1}{k^2} |R_o(kN)| , \qquad (7)$$

$$k = 1, 3, 5 - - -$$

where $|R_o(kN)|$ represents the values of the sensor's MTF at frequencies kN where kN are the odd harmonics of frequency N. For the vidicon used in our experiments, the square wave amplitude, modulation transfer and square wave flux responses are as shown in Fig. 5.

Using the square wave flux response concept, we calculate the $SNR_{DI}$ from the relation

$$SNR_{DI} = \left[\frac{2tn_v \Delta f_v}{\alpha}\right]^{\frac{1}{2}} \frac{R_{SF}(N)}{N} \left(\frac{\Delta i_o}{I_n}\right) . \qquad (8)$$

In the above, $\Delta i_o$ is the peak-to-peak signal current for a broad area pattern (unity modulation transfer function) and $I_n$ is the rms noise that is added to the camera

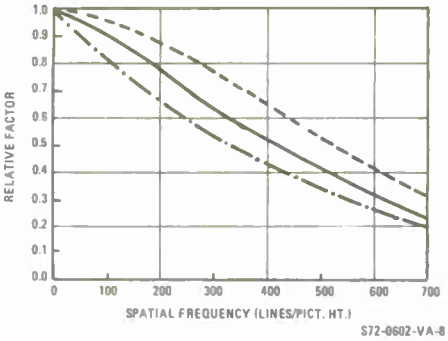



Fig. 5 Square Wave Amplitude (- - -), Modulation Transfer (———), and Square Wave Flux (- · -) Functions for the 1½" Vidicon used in the Bar Pattern, Recognition and Identification Experiments.

generated image. Real cameras, of course, have a response that is a function of frequency and the value of $\Delta i_o$ in the video channel for square wave inputs becomes $\Delta i_{p-p}$, the peak-to-peak value of the video signal when the frequency effects are included, that is, we have

$$\Delta i_{p-p} = \Delta i_o R_{SQ}(N) \qquad (9)$$

and

$$SNR_{DI} = \left[\frac{2tn_v \Delta f_v}{\alpha}\right]^{\frac{1}{2}} \left(\frac{1}{N}\right) \frac{R_{SF}(N)}{R_{SQ}(N)} \frac{\Delta i_{p-p}}{I_n} \qquad (10)$$

where $R_{SF}(N)$ is the value of the flux factor at N, $R_{SQ}(N)$ is the value of the square wave response at N and $\Delta i_{p-p}$ the value of the peak-to-peak signal corresponding to N as measured in the output of he video channel. For calculation purposes, t, the integration time of the eye is taken to be 0.1 sec and $\alpha$, the picture aspect ratio is 4/3. At low spatial frequencies the displayed images approach a square wave while at high spatial frequencies, above about 500 lines/picture height, the displayed images were nearly pure sine waves. The main point of the above discussion is to note that the $SNR_{DI}$ are calculated on the basis of mean signal to rms noise since we have corrected the measured signal using the square wave amplitude and square wave flux factors.

For the first series of experiments, the test images were bar patterns of

various spatial frequencies and 5:1 height-to-width ratio. The observer was required to state whether or not the pattern displayed was resolvable as the images $SNR_{DI}$ were randomly varied. Chance was not involved since the patterns were always present on the display. A number of bar patterns of various spatial frequencies were displayed at the same time and the observer was asked to indicate the pattern of highest frequency that was barely visible as the signal-to-noise ratio was systematically varied from high levels to low levels and the reverse in steps of 1 db. The fraction of patterns resolved by this "method of limits" are plotted in Fig. 6 as a

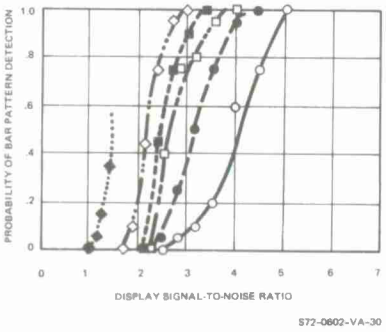



Fig. 6 Probability of Bar Pattern Detection vs $SNR_{DI}$ for Bar Patterns of Spatial Frequency ○ 104, ● 200, □ 329, ■ 396, ◇ 482, and ◆ 635 Lines per Picture Height. Bar Height-to-Width Ratio was 5 in All Cases.

function of display signal-to-noise ratio and the thresholds for 50% probability of detection are plotted in Fig. 7.

To test the notion that the eye integrates over the total length of a bar, we employed bar patterns of various length-to-width ratios. A typical result is shown in Fig. 8 for a 396 line bar pattern. The thresholds for three different bar height-to-width ratios are shown in Fig. 9 and are seen to be similar to those shown in Fig. 7. These data were obtained using random signal-to-noise ratios rather than the "method of limits" used in the first series of measurements. Either method gives about the same results. Since the "method of limits" is most often used in

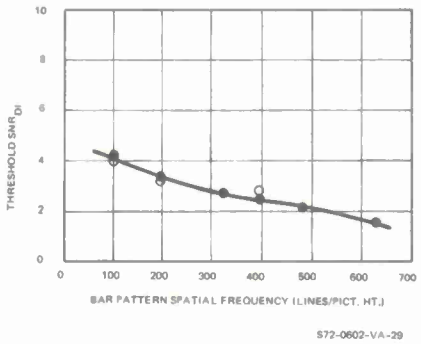



Fig. 7 Threshold $SNR_{DI}$ vs Bar Pattern Spatial Frequency Obtained Using Two Different Experimental Techniques, ● Method of Limits and ○ Methods of Random SNR Variation.

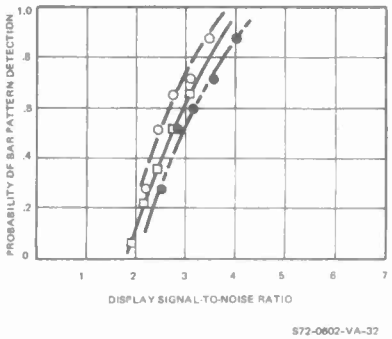



Fig. 8 Probability of Bar Pattern Detection vs $SNR_{DI}$ for a 396 Line Bar Pattern of Bar Length-to-Width Ratio □ 5:1, ○ 10:1, and ● 20:1.

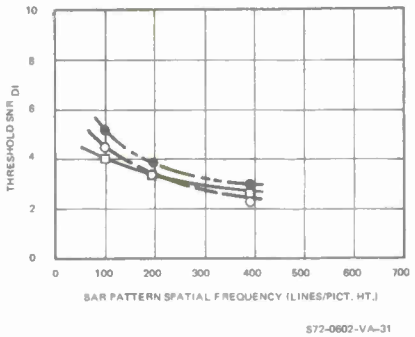



Fig. 9 Threshold $SNR_{DI}$ vs Bar Pattern Spatial Frequency for Three Bar Length-to-Width Ratios of □ 5:1, ○ 10:1, and ● 20:1.

lab evaluations of TV equipments, it is interesting to note that the method is appropriate. Further efforts to confirm the utility of the "method of limits" were made with the results shown in Fig. 7.

In the above experiments, the bar patterns were vertically oriented with their longitudinal axes perpendicular to the direction of scan. The results with the axes horizontal are shown in Fig. 10. One observer was used for the

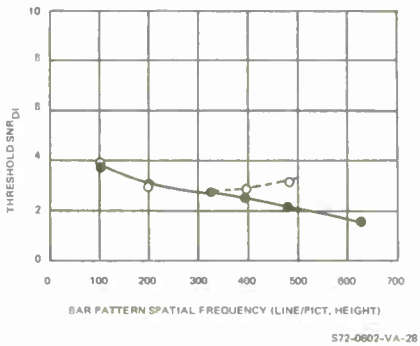



Fig. 10 Threshold $SNR_{DI}$ Required to Recognize the Presence of • Horizontally Oriented and o Vertically Oriented Bar Pattern vs Bar Pattern Spatial Frequency.

experiment. Also shown in Fig. 10 are the results for vertical bars for the same observer (the same results as those shown in Fig. 7). As can be seen, the orientation is immaterial at the low spatial frequencies but the thresholds for the high spatial frequencies increase with horizontal bars. Indeed, the 635 line pattern could only be seen with noise added to the signal.

In the next bar pattern experiment, viewing distance was varied in discrete steps from 14" to 28" to 56". The bars were vertically oriented. The results are as shown in Fig. 11. At the short distances, the low pattern frequencies become less detectable while the reverse was true at the long viewing distances. It is clear from Fig. 11 that, for a given line number, there is a viewing distance that minimizes threshold $SNR_{DI}$.

The above experiments involved 5 observers on the average and 500 to 2000 observations per experiment.

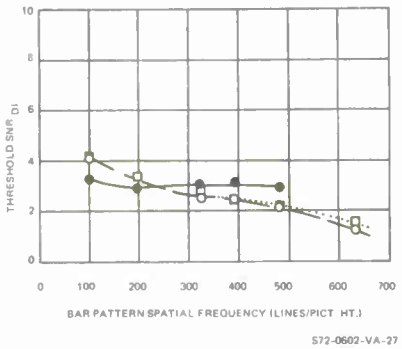



Fig. 11 Threshold $SNR_{DI}$ vs Bar Pattern Spatial Frequency for Display-to-Observer Viewing Distances of o 14", □ 28" and • 56".

## LEVELS OF DISCRIMINATION

The purpose of many electro-optical equipments is to augment an observer so that he can discern a scene object at longer range than he could with his unaided eye. These objects may be seen with greater or lesser clarity depending on the object's size, distance, contrast, radiance, and the sensor's resolving power. Here, we are concerned with a condition where a scene object needs to be discerned with sufficient clarity to serve some intended function and, that the observer is highly motivated to do so. The words "sufficient clarity" should be stressed. In one case, it may be sufficient to merely detect a blob such as a channel buoy while in other cases, much higher acuity is needed. For example, it is of no use to televise and record a burgulary if the recording's acuity is insufficient to identify the burgular in a court of law.

It is intuitively obvious that an object may be seen with greater or lesser clarity depending on the object's size, distance, contrast, brightness and the capability of the observer alone or the observer as augmented by an imaging device. In planning a surveillance task, the maximum or threshold range at which a given object can be barely discerned at the desired level of discrimination is clearly of interest. At very long range, a scene object may appear first as a blob. Moving closer, the object can be discerned to have some shape such as a rectangle.

Closer yet, the object becomes able to classify the object as to type and finally, to identify the object positively. (Johnson, 1958) has arbitrarily divided these levels of object discrimination into 4 categories which are:

Table I
Levels of Object Discrimination

| Classification of Discrimination Level | Meaning |
|---|---|
| Detection | An object is present. |
| Orientation | The object is approximately symmetrical or unsymmetrical and its orientation may be discerned. |
| Recognition | The class to which the object belongs may be discerned (e.g., house, truck, man, etc.). |
| Identification | The target can be described to the limit of the observer's knowledge (e.g., motel, pickup truck, policemen, etc.). |

It is readily evident that higher resolution is needed to identify an object than to just detect it. To obtain a quantitative feel for the problem, Johnson performed a series of experiments using electro-optical sensors. In these experiments, an attempt was made to correlate the detectability of a bar pattern of a given spatial frequency with the level of object discrimination. The procedure was to increase the object range until it was just barely detected (or recognized, etc.). Then, a bar pattern was placed in the field of view and its spatial frequency was increased until it could barely be resolved at the same range. The spatial frequency of the pattern was specified in terms of the number of lines in the pattern subtended by the object's minimum dimension as illustrated in Fig. 12 where the object, for the recognition case, subtends 8 lines.

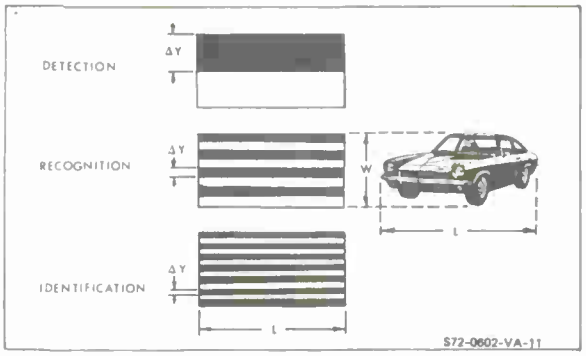



Fig. 12 Resolution Required per Minimum Object Dimension to Achieve a Given Level of Object Discrimination Expressed in Terms of an Equivalent Bar Pattern.

Johnson's results, as tabulated below are not unexpected. If the observer could only just resolve a coarse pattern corresponding to 2 bars per minimum object dimension, the level of object discrimination was limited to detection. With higher acuity, a bar pattern of higher spatial frequency could be discerned and the level of object discrimination increased in turn.

Table II
Johnson's Criteria for the
Resolution Required per Minimum
Object Dimension vs Discrimination Level

| Discrimination Level | Resolution per Minimum Object Dimension (TV Lines) |
|---|---|
| Detection | $2.0 \begin{smallmatrix} +1.0 \\ -0.5 \end{smallmatrix}$ |
| Orientation | $2.8 \begin{smallmatrix} +0.8 \\ -0.4 \end{smallmatrix}$ |
| Recognition | $8.0 \begin{smallmatrix} +1.6 \\ -0.4 \end{smallmatrix}$ |
| Identification | $12.8 \begin{smallmatrix} +3.2 \\ -2.8 \end{smallmatrix}$ |

This table has been widely used and misused by systems designers from the time of its publication to the present. The misuse stems from the neglect of additional requirements imposed by Johnson, to wit, that the "signal-to-noise" ratio and image contrast must also be sufficient. However, it was not too clear how these quantities were to be measured and

calculated and thus, the further requirements were neglected in many cases. However, many competent designers did use the sensor's threshold resolution vs scene irradiance curves in estimating the level of discrimination. Since the threshold curves do contain image signal-to-noise as a factor in their measurement, estimates made on this basis turn out to be reasonable if not precise.

Most sensors are characterized by an absolute limiting resolution. If the sensor sensitivity at a given scene irradiance level and object contrast is sufficient to realize the limiting resolution but the resolution is still not sufficient to perform the desired discrimination task, further increases in scene irradiance level will be to no avail. The only solution is to move closer. On the other hand, if the resolution were sufficient at a given irradiance level, a decrease in scene irradiance level could cause the acuity of a sensor and observer combination to fall below the level required for a wanted level of object discrimination.

To illustrate, suppose it is desired to discriminate the vehicle of Fig. 13.

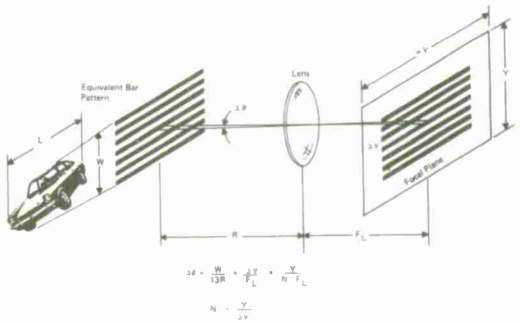



Fig. 13 Object and Image Resolution Required to Identify an Object in Terms of the Equivalent Bar Pattern and the Scene, Lens and Focal Plane Parameters.

The minimum scene resolution required for a vehicle of minimum dimension w ft will be $w/k_d$ ft where $k_d$ is the number of lines needed for the wanted level of discrimination per Table II. At range, R, the sensor-augmented-observer's angular resolution, $\Delta\theta$, must then be $w/k_dR$ radians. The image size, $\Delta y$, will be equal to $F_L \cdot \Delta\theta$ where $F_L$ is the lens focal length. Suppose the primary sensor, upon which the scene is imaged is a television camera tube.

The resolution of the TV camera tube is measured by placing a bar pattern of spatial frequency N lines/picture height into the field of view and determining whether the test bar pattern can be resolved or not. If the number of lines N, which the camera can resolve is equal to or greater than the reciprocal dimensions N of the object image, the object can presumably be resolved as well. For this comparison to be valid, the test bar pattern must meet certain other conditions as will be discussed.

As a numerical example, suppose the vehicle of Fig. 13 is of height 5' and is to be identified so that $k_d = 13$, the scene resolution must be about $4\frac{1}{2}$" corresponding to an angular resolution $\Delta\theta$ of $64 \times 10^{-6}$ radians at a range of 6000 ft. For a typical TV camera tube picture height of 15 mm, and a threshold resolution of 250 lines/pict. ht., a lens focal length, $F_L$, of 940 mm or 37" will be needed.

Johnson's equivalent bar pattern results were held to be valid if the bar spacing is chosen on the basis of Table II, if the bar pattern contrast is the same and, if the image signal-to-noise ratio, or SNR, is sufficient. It might then be inferred that image contrast, resolution and SNR are independent quantities. However, we shall derive functional relationships between them. Then Johnson's requirements will reduce to one;* namely, that an object should be discriminated at the desired level if its signal-to-noise ratio at the output of the observer's retina after processing and interpretation by the brain is large enough. Obviously, the signal-to-noise ratio as defined in this manner, is not directly measurable but can be indirectly measured through psychophysical experimentation.

To show a typical result, we will describe a recognition experiment. Four

* Using the equivalent bar pattern concept to define resolution.

vehicles, a tank, a van truck, a half track with antenna and a derrick bulldozer were televised. Noise of known amount was added to the imagery and the probability of the observer correctly recognizing the type of vehicle was determined as a function of the vehicle image's SNR with the result shown in Fig. 14. The method of determining the SNR will be discussed in detail below. The main point is that it appears to be possible to associate a signal-to-noise ratio with an image of a real object based on its dimensions, contrast and the sensor parameters. We see in Fig. 14, that the $SNR_{DI}$ required

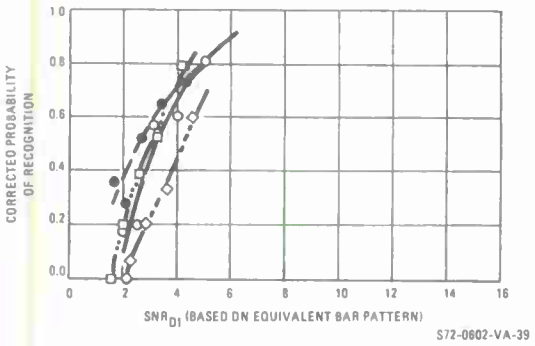



Fig. 14 Probability of Recognition vs $SNR_{DI}$ for a O Tank, ◇ Radar Half Track, □ Van Truck, and ● Derrick Bulldozer. Background was Uniform.

to recognize the vehicles with 50% probability is about 3. To test Johnson's approach, the vehicles were replaced by equivalent bar patterns of spacing 1/8 the minimum vehicle dimension and with individual bars of length equal to the vehicles length. The bar pattern imagery was mixed with noise and probability of bar pattern detection was determined as a function of image's SNR with the result shown in Fig. 15. For 50% probability of detection, an $SNR_{DI}$ of 2.8 is seen to be needed. The correlation between the recognition of the real object and the detection of the equivalent bar pattern is seen to be very good -- and perhaps, better than is warranted when other factors are considered as we shall discuss.

The above experiments were conducted using a uniform white background. The vehicle areas were approximately 0.057 $in^2$ on the 8" x 10.7" display and subtended angles of about $0.34^o$ by $0.68^o$ at the observer's eye. The vehicle types and video SNR were randomly varied; and the probabilities of recognition, corrected for chance, were determined. The $SNR_{DI}$'s for the various images were calculated on the basis of the area of a bar whose length and width are equal to the length of the vehicle's image and the width of the vehicle's image divided by 8. This is in accord with the equivalent bar pattern concept discussed above and illustrated in Fig. 12. We note, however, one difference between the calculations for the bar pattern and the vehicular image's $SNR_{DI}$. In the case of the vehicular image, the signal amplitude was measured from the background signal level which was approximately constant, to the peak object signal level. For the "equivalent bar patterns", the signal levels were measured in terms of the mean signal excursion within the bar pattern area in the periodic direction. Had the peak-to-peak excursions about the average signal within the vehicle area been used (when the object is imaged against a uniform background), the threshold $SNR_{DI}$ would have been somewhat lower.

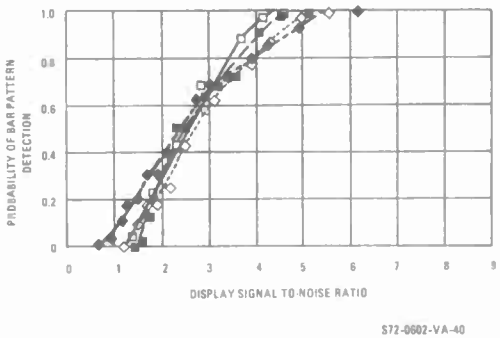



Fig. 15 Probability of Bar Pattern Detection vs $SNR_{DI}$ for Pattern Area Equal to Average Real Object Area of N = □ 329, 7 Bars; ■ 396, 7 Bars; ◇ 482, 9 Bars; ◆ 635, 11 Bars.

These difficulties result from the necessity of defining an image area and a signal excursion in order to calculate an $SNR_{DI}$ threshold. In this connection, we observe that the criterion for bar pattern "recognition" is that the observer must be able to discern a modulation within the bar pattern whereas for vehicle image

recognition, the vehicle's outline must be discerned. This outline may have periodic features but is more likely to be aperiodic. While we recognize and even emphasize these differences, we do not now have better criteria to suggest. It is felt that recognition predictions based on the equivalent bar pattern concept will tend to be somewhat pessimistic because the effect of an MTF is generally more severe on periodic objects than on aperiodic objects. It is interesting to observe that the thresholds measured for recognition of various vehicle types are nearly constant but that the scatter in the data increases as we move from the lowest level of discrimination (detection) to the highest (identification). An alternative concept for recognition and identification might be to define some sub-area on the image and treat it as an aperiodic object. Intuitively, we would expect that range predictions made on this basis would be optimistic because sensor MTF's will tend to reduce the modulation of image detail features which are not well separated from one another.

Each object used in the recognition experiments did have a feature which approached the idea of an isolated object feature such as a boom arm, a turret, etc. It may be that a single feature is characteristic of a specific object and to recognize the object, the characteristic feature must be discerned. In each case, the maximum signal excursion in the "characteristic feature" was the same as the largest found for the whole target, or nearly so. Also, the width of the feature was about 1/8 the total minimum width of the object. A similar situation was found in the identification experiments where the features needed were closer to 1/13 the object's minimum width.

In Fig. 16, we show the signal excursions for the tracked bulldozer on selected horizontal lines from the top to the bottom of the vehicle with line 1 being just above the object and line 17 just below. These traces were taken on every other line from a line selector oscilloscope. The dominant features of the bulldozer can be located in the traces of Fig. 16. The boom arm is in traces 3 through 6, the cab in 6 through

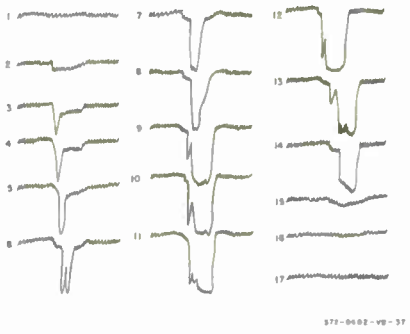


Fig. 16 Waveform of Derrick Half Track Along the Horizontal as a Function of Vertical Position in Scan Lines.

8 and the bulldozer blade in 9 through 13. The boom appeared to be the most characteristic feature and the width was of the order of 1/8 width of the bulldozer.

The second recognition experiment was performed by superimposing a background transparency over the same vehicle transparencies as were used above. The background transparency has regions containing a road, a grassy field, and grass among trees. In the experiment, a vehicle was randomly chosen and located within one of the randomly chosen background regions noted above. In all, 12 vehicle-background combinations were shown at 5 randomly selected $SNR_{DI}$ values. The $SNR_{DI}$ was calculated as in the first recognition experiment. The results using a road background are shown in Fig. 17. The spread in the data was small and a threshold value of $SNR_{DI}$ is 3.8 which is only 15% higher than that noted for a uniform background. The results with grass background are shown in Fig. 18 and give a threshold value of $SNR_{DI}$ of 4.1; a value that is 24% higher than for a uniform background. In Fig. 19, a threshold value of 5.0 was obtained for the recognition of vehicles in a grass and tree background which is 52% higher than the uniform background case and 31% higher than the road background case and 22% higher than the grass background case. As one would expect, increases in background complexity increase the $SNR_{DI}$ thresholds observed but the increases are not large.

Considering Figs. 14, 17, 18 and 19, it is seen that, on the average, the derrick bulldozer and van truck have an $SNR_{DI}$ threshold that is about 26% lower than that required for the tank and radar half track. By visually comparing these objects, it appeared reasonable that less resolution would be needed to recognize the derrick bulldozer and van truck. Assume that, in general, a different resolution is required to recognize each object and that if this resolution were used in the $SNR_D$ calculation (instead of 8)

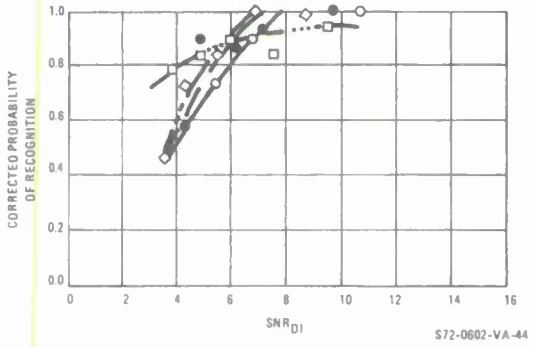



Fig. 17 Probability of Recognition vs $SNR_{DI}$ for a O Tank, ◇ Radar Half Track, □ Van Truck, and ● Derrick Bulldozer Against a Road Background.

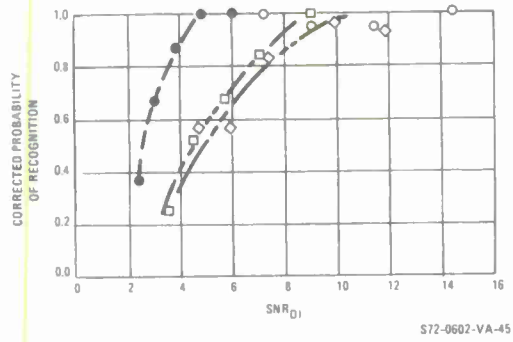



Fig. 18 Probability of Recognition vs $SNR_{DI}$ for a O Tank, ◇ Radar Half Track, □ Van Truck, and ● Derrick Bulldozer Against a Grass Background.

that the value of $SNR_D$ would be the same for the various images. The spread in the measured data, using the value of 8, should correspond to the spread in the required resolution and, from the spread we have that the required resolution for target recognition is $8 \pm 2$ lines.

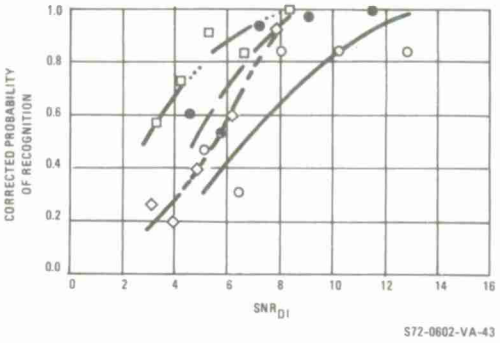



Fig. 19 Probability of Recognition vs $SNR_{DI}$ for a O Tank, ◇ Radar Half Track, □ Van Truck, and ● Derrick Bulldozer Against a Grass-Trees Background.

Next, we progressed to vehicle identification experiments. The images for this experiment were transparencies of the M47 Patton, the M48 Centurion, the Panther and the Stalin tanks as shown in Fig. 20. As for the previous vehicle recognition case, the photographs were taken to show the sides and tops of the vehicles from an angle of $45^o$ from the horizontal. Two image sizes were used. In the first experiment, the average

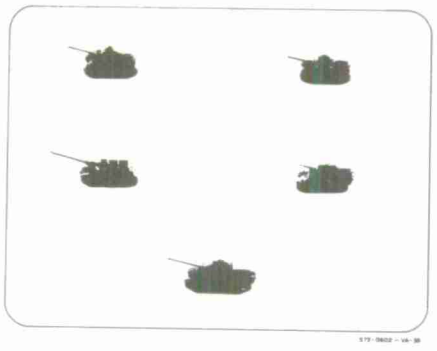


Fig. 20 Photographs of Tank Models Used in Identification Experiments. Upper Left - M47, Upper Right - M48, Center Left - Stalin, Center Right - Panther, Lower Center - Centurion.

image on the 8" x 10.7" display was 0.9 $in^2$ and subtended $1.3^o$ x $2.6^o$ at the observer's eye; and in the second experiment, the image was about 2.2 $in^2$ and the angular extent was about $2^o$ x $4^o$.

The objects and SNR values were randomly varied and the probabilities of correct identification were corrected for chance. For the calculation of $SNR_{DI}$, the image area was assumed to be a bar whose length and width are equal to the length of the vehicles image and the width of the vehicle's image divided by 13 using the equivalent bar pattern concept discussed and illustrated in connection with Fig. 12. The results of the first identification experiment are shown in Fig. 21 where

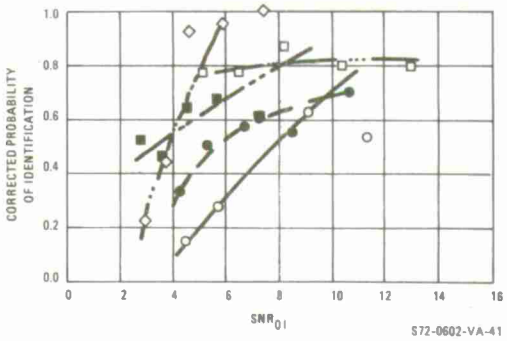



Fig. 21 Probability of Identification vs $SNR_{DI}$ for Tanks of Angular Extent $1.3^{o}$ x $2.6^{o}$ Relative to the Eye Against a Uniform Background. Tanks were O M47, □ Centurion, ■ Panther, and ◇ Stalin.

we obtain a threshold $SNR_{DI}$ of 5.2. This is nearly identical to the value obtained in a separate experiment for an equivalent bar pattern which had 13 bars with a length-to-width ratio of 26 and a spatial frequency of 137 lines/pict. ht.

In the second identification experiment, we obtained a threshold value of 6.8 as shown in Fig. 22. An estimate of 6.2 was obtained for an equivalent bar pattern of 26:1 length-to-width ratio and a spatial frequency of 101 lines/pict. ht. From the Figs. 21 and 22, it is seen that, on the average, the Panther and Stalin tanks are seen with an $SNR_{DI}$, at threshold, which is about 42% smaller than that required for the M48 and M47. From the photographs of the tanks in Fig. 20, it is seen that the Stalin and Panther tanks appear more distinctive than the M47 and M48 so that the lower threshold $SNR_{DI}$ is not entirely unexpected.

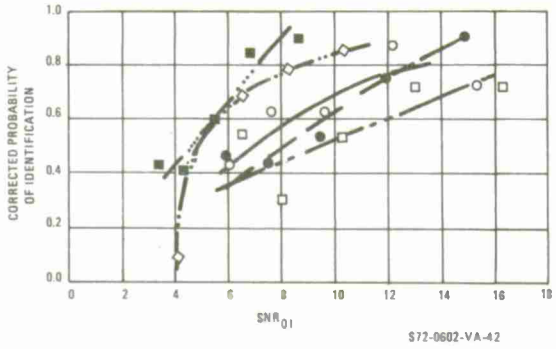



Fig. 22 Probability of Identification vs $SNR_{DI}$ for Tanks of Angular Extent $2^{o}$ x $4^{o}$ Relative to the Eye Against a Uniform Background. Tanks were O M47, □ Centurion, ■ Panther, and ◇ Stalin.

In summary, it was hypothesized that the recognizability and identifiability of "real world" objects could be correlated with the discernability of an "equivalent bar pattern." The "equivalent bar pattern" was defined (in image space) in terms of a single bar in the pattern of length equal to the length of the image of the "real world" object and width equal to the real world image's width divided by $k_d$. The factor, $k_d$, is a number associated with a given discrimination level, e.g., 8 for recognition. When the bar pattern is defined in the above manner, it was found experimentally that the $SNR_{DI}$ required to liminally discern the equivalent bar pattern was very nearly equal to the $SNR_{DI}$ required to recognize the "real world" image when its $SNR_{DI}$ is calculated on the basis of its area divided by $k_d$. It might be thought that defining the area of the bar in the equivalent bar pattern as 1/8 the area of the "real world" image's area and then calculating the $SNR_{DI}$ for the real world image on the basis of 1/8 its area is redundant but this is not so. The key reason for defining the bar pattern in the manner described is to define the required level of sensor resolution in terms of a spatial frequency. This permits us to take into account the effects of finite sensor apertures. As we noted previously, Johnson specified that an image, to be discerned at a given level of discrimination must be both resolved and have a sufficient SNR. In the above formulation, using the equivalent bar

pattern approach, $SNR_{DI}$ is a function of resolution so that if the $SNR_{DI}$ is sufficient, the resolution will be sufficient as well. The threshold $SNR_{DI}$'s required for the various levels of discrimination are estimated in Table III. These thresholds apply to a 50% probability but can be adjusted for any other level of probability by use of Fig. 23.

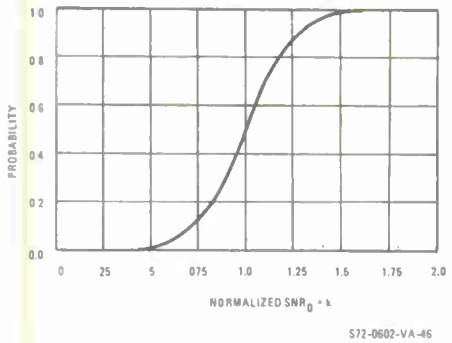



Fig. 23 Probability vs Normalized $SNR_{DI}$. For Any Probability Value, Obtain $SNR_{DI}$ From Table III for 50% Probability. Find Value of k for Desired Probability and Multiply Value of $SNR_{DI}$ by k to Obtain New Value of $SNR_{DI}$ Required.

We noted an important caution in connection with the method of calculation. The "real world" objects are not periodic in general but rather, assemblages of aperiodic objects of different sizes. These aperiodic objects which make up the object may be relatively isolated as in the case of the derrick on the tracked bulldozer or more clutter-like as in the case of the half track. Thus, the equivalent bar pattern approach weights cluttered scenes more heavily and may tend to be somewhat pessimistic, particularly at high line numbers.

## References


Barnes, R., and Czerny, M., Z Phyzic, 79, 1932

de Vries, H. L., Physics X, No. 7, July 1943

Johnson, J., Image Intensifier Symposium, Ft. Belvoir, Va., AD220-160, Oct. 1958

Rose, A., J. Opt. Soc. Am., 38, 196, 1948

Rosell, F. A., J. Opt. Soc. Am., 59, 539, 1969

Rosell, F. A., and Willson, R. H., AFAL-TR-71-137, Wright Patterson AFB, Dayton, Ohio, May 1971

Rosell, F. A., and Willson, R. H., AFAL-TR-72-279, Wright Patterson AFB, Dayton, Ohio, August 1972

Schade, O. H., Sr., RCA Review, Sept. 1967


TABLE III Best Estimate of Threshold $SNR_{DI}$ for Detection, Recognition and Identification of Images.

| Discrimination Level | Background | $k_d$ TV Lines per Minimum Dimension | Threshold $SNR_{DI}$ for Spatial Frequency (Lines/Pict. Ht.) of | | | |
|---|---|---|---|---|---|---|
| | | | 100 | 300 | 500 | 700 |
| Detection | Uniform* | 1 | ← 2.8 → | | | |
| Detection | Clutter | 2 | 4.8 | 2.9 | 2.5 | 2.5 |
| Recognition | Uniform | 8 | 4.8 | 2.9 | 2.5 | 2.5 |
| Recognition | Clutter | 8 | 6.4 | 3.9 | 3.4 | 3.4 |
| Identification | Uniform | 13 | 5.8 | 3.6 | 3.0 | 3.0 |

* Treated as an Aperiodic Object.

# THE PREDICTION OF SIGNAL STRENGTH REQUIRED FOR IMAGE DETECTION/RECOGNITION ON A RASTER GENERATED DISPLAY

by

Robert Sendall
Electro-Optical Systems
Pasadena, California 91107

OFFICE OF NAVAL RESEARCH
TARGET ACQUISITION SYMPOSIUM
NAVAL TRAINING CENTER, ORLANDO, FLORIDA / 14,15,16 NOVEMBER 1972

# THE PREDICTION OF SIGNAL STRENGTH REQUIRED FOR IMAGE DETECTION/RECOGNITION ON A RASTER GENERATED DISPLAY

by

Robert Sendall
Electro-Optical Systems
Pasadena, California 91107

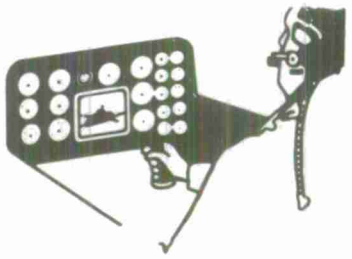
## ABSTRACT

The theory of detection of visual anomolies on a raster generated display is independent of the type of sensor supplying the information. While recognition and therefore detection in clutter requires training, the basic theory is independent of the sensor. This paper presents an analytical approach which can be used with any sensor. It is compatible with the famous J. Johnson criteria and is based upon the theory that the operator is a spatial and temperal integrator making detection decisions dependent upon the perceived signal-to-noise ratio.

The paper shows effects of system component modulation transfer functions as they affect both signal and noise. Both white and colored displayed noise will be discussed reflecting the effects of various possible noise source injection points. In general the perceived signal-to-noise ratio will be related to the video signal-to-noise ratio and the probability of detection will be related to the perceived signal-to-noise ratio.

Many discussions with F. Rosell and a few with J. M. Lloyd have aided in the development of the analysis which is essentially an adaptation of the papers of O. Schade.

COMPLETE PAPER HAS BEEN PUBLISHED UNDER AFAL-TR-72-374 "E/O SENSOR PERFORMANCE ANALYSIS AND SYNTHESIS (TV/IR COMPARISON STUDY)" - ROBERT SENDALL AND FREDERICK ROSELL - AUTHORS



# DETECTABILITY THRESHOLDS FOR LINE-SCAN DISPLAYS

by
Robin L. Keesee
Virginia Polytechnic Institute and State University
Blacksburg, Virginia

OFFICE OF NAVAL RESEARCH
TARGET ACQUISITION SYMPOSIUM
NAVAL TRAINING CENTER, ORLANDO, FLORIDA / 14,15,16 NOVEMBER 1972

# DETECTABILITY THRESHOLDS FOR LINE-SCAN DISPLAYS

Robin L. Keesee
Virginia Polytechnic Institute and State University
Blacksburg, Virginia

## ABSTRACT

To evaluate the validity of the Modulation Transfer Function Area (MTFA) as a proposed unitary measure of TV image quality, the visual thresholds for detection of tribar targets on a video monitor as functions of target modulation, target spatial frequency, white noise level, and noise passband is determined for three representative TV systems having different line rates and bandwidths. For these three specific systems, optical measurements of the displayed video noise measured in luminance terms, will be compared with the electrical rms measurements of noise inserted into the video system. Implications of these types of measurements and their importance in specifying the displayed image quality will be stressed.

## INTRODUCTION

In the ongoing search for a summary measure of TV image quality, the modulation transfer function area (MTFA) has become increasingly attractive to some researchers (Snyder, 1970). Charman and Olin (1965) performed an analytical validation of the concept for photographic systems, and a group at Boeing (Borough, Fallis, Warnock, and Britt, 1967) showed that there was a high correlation between MTFA and photo-interpreters' opinion of photographic image quality. Although preliminary results indicated that the MTFA is a valid measure of TV image quality (Snyder, 1972), a thorough evaluation and practical utilization of the MTFA as a video image quality metric has been impeded by the lack of satisfactory detectability threshold data.

Conventional contrast thresholds are unsuitable because they do not include the effects of the sampling process of the video raster. Coltman and Anderson (1960) reported the experimental determination of some detectability thresholds as verification of a theoretical derivation. The usefulness of these curves is perhaps suggested by quoting from their paper:

> "...The amount of data taken was limited, and conditions of surround brightness, time interval between tests, etc., were not carefully controlled, so that the data presented here do not constitute a definitive study of this particular visual parameter...."

Clearly, an experimental determination of line-scan detectability thresholds is needed. This paper describes an attempt at filling this need.

The quality of the MTFA evaluation and the extent of the use of the MTFA as an image quality metric depends on the generality of the detectability threshold curves. There are several things to be considered in planning for this generality. The most obvious element in such planning is that the generality is directly proportional to the number of different detectability thresholds determined. Thus, the threshold curves were found for several different system MTF's and for a full range of spatial frequencies, modulations, and noise levels. The effect of noise passband on the detectability threshold curves was also investigated. Generality is especially enhanced by describing quantitatively the input to the TV system and the system output. Further, the experiment

was carefully controlled so that the input/output quantification remained appropriate throughout the experiment.

The psychophysical method of adjustments was used to reduce subject error and to help eliminate experimenter bias. The method of adjustments required the dependent variable to be continuous. Among system bandwidth and line rate, target size, target modulation, noise level, and noise passband, only noise level could conveniently be made a continuous dependent variable. For data stability, a large number of data points for each combination of values of the independent variables was needed, which dictates the use of a few highly practiced subjects and many trials. The input/output of the video system.was measured including camera lens, camera, etc., so that the resulting methods of measurement are generally applicable to other systems. Using the complete video chain precluded any electrical generation of targets. Further, since the stimulus to the subject in finding the detectability thresholds is light emitted from the front of the monitor, then the video output measurements should be photometric.

## DETERMINATION OF DETECTABILITY THRESHOLDS

### Design of the Experiment

A summary of the specific experimental design is given in Figure 1. Three systems with different MTF's were used. Three different MTF's were obtained by operating the closed-circuit television (CCTV) at 32 MHz with a 1225-lines per frame rate, at 16 MHz bandwidth and 945 lines, and at 8 MHz bandwidth, 525 lines. The square wave responses corresponding to these three MTF's are given in Figure 2. Three or four different noise passbands at each bandwidth/line rate combination were included, as shown in Table 1. The targets were a series of 8 x 10-inch photographs on the 1951 USAF tribar pattern with black or gray bars against a white background. The bar or square-wave target was chosen for simplicity of definition of the subject response criterion and for ease of displayed target measurement. The targets were made in seven spatial frequencies with eight modulations at each spatial frequency and were displayed with the bars perpendicular to the raster lines. The raster was vertical, as viewed by the subject.

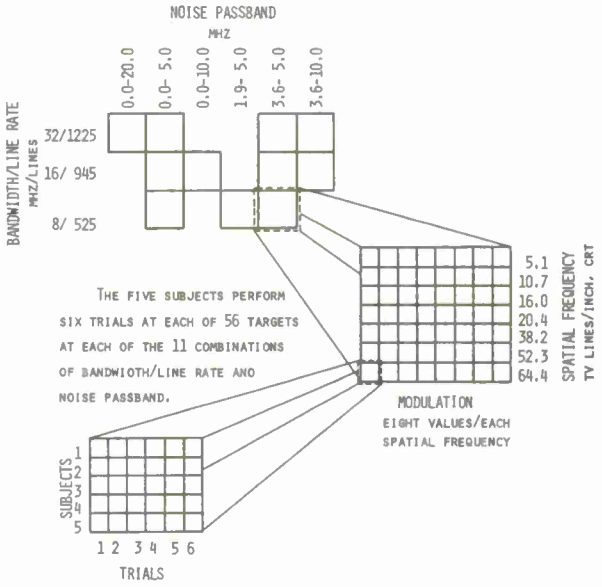



Figure 1. Design of the Experiment

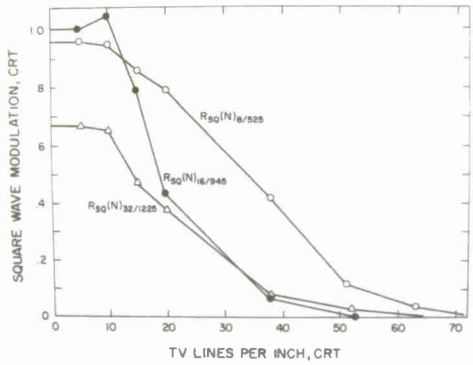



Figure 2. Square-Wave Response for the Three Systems

Subjects were three male and two female paid students having a minimum of 20/22 binocular and 20/25 monocular visual acuity without correction and without any other visual anomalies as tested with the Bausch & Lomb Otho-rater. Each subject received six trials at each combination of spatial frequency, modulation, noise passband, and bandwidth/line rate, for a total across all subjects of 18,480 data points.

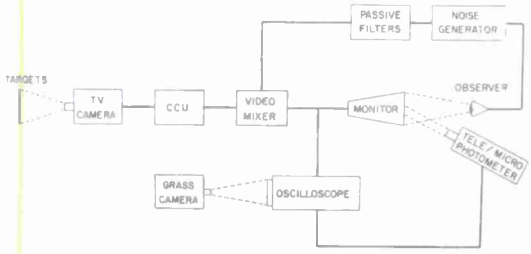



Figure 3. Video System

The television system is block diagrammed in Figure 3. One of the fifty-six targets was placed before the Cohu 6100 camera. Composite video from the Cohu 6900 camera control unit (CCU) passed through a 35 MHz mixer to a 17-inch Conrac RQA monitor. A General Radio 1383 Random Noise Generator produced a white noise in the 20 Hz to 20 MHz spectrum. Narrower noise passbands were achieved by passing this noise through one or two passive filters. Bandwidth and line rate changes were made by changing components within the camera and the CCU.

Components of a Gamma Scientific photometer were assembled in either a telephotometer or microphotometer mode to make measurements at the CRT monitor. The subject, monitor, noise generator, noise filters, photometer, and the first experimenter were in one room. In an adjoining room were the camera, CCU, targets, and the second experimenter. An intercom connected the two experimenters. Both of these rooms were climate controlled.

Each subject was seated so that the eye-to-CRT distance was 40 inches and the eyes were level with the center of the CRT. The noise generator was placed so that the noise level could be adjusted by the subject while the rms noise level meter was visible only to the experimenter. The experimental room was dark except for the monitor and a small lamp for the experimenter. No light from this small lamp fell in the subject's field of view.

## Procedure

The normal daily experimental procedure began with allowing one hour of warm-up for all of the electronics prior to the start of data collection. At the end of this warm-up, the system was checked for calibration using a gray scale target in front of the camera. Overall video and blanking levels at the CCU were checked and adjusted to constant levels. Composite video from the CCU was monitored on an oscilloscope and any changes necessary in the CCU adjustments were made to return the video levels for each gray bar to set values. With the electrical input to the monitor thus normalized, the image on the monitor of the gray scale target was viewed with the telephotometer. The luminance of three particular gray bars, one near white, the second middle gray, and a third near black, was measured, and the contrast and brightness controls of the monitor were adjusted to bring the luminance of these bars within certain tolerances. This calibration procedure was performed at the beginning of each experimental session and repeated every hour during the session. As an example of the importance of these procedures, the luminance of the near white gray bar was adjusted during calibration to be between 18.0 and 18.5 footlamberts. After an hour of operation, the luminance would drift one or two footlamberts. If this drift had continued unchecked, the results would have been suspect since the stimuli were truly random. There is no known evidence to suggest that this drift is peculiar to this TV system.

After calibration, a subject was seated and the seat height was adjusted so that the subject's eyes were in proper position and the subject was comfortable. The first experimenter requested a target at random. The second experimenter placed the requested target before the camera. The subject increased the noise level until he could no longer determine that there were three separate bars. The first experimenter recorded this noise level, increased the noise level until well past the point where the target was not visible, and told the subject to proceed. The subject then reduced the noise level until he could just determine that there were three separate bars. This noise level was then recorded, completing the pair of trials. The subjects' criterion was not that of a detection task or a recognition task; the criterion was simply the existence or non-existence of three, separate bars.

In theory, since the ascending trial gave a noise level slightly higher than the actual threshold noise level and the descending trial gave a noise level slightly below the threshold, the mean of these two trials would be the actual threshold. With practice, a pair of trials was completed in about 12 seconds. Subjects worked for periods of 20 to 25 minutes and were then given a five-to-ten-minute break. Three of these periods plus the calibration procedures filled a two-hour experimental session. Each of the five subjects worked each day during the experiment and for no more than two hours. This work was demanding and any more than two hours per day per subject caused subject fatigue and erratic performance.

## Results

Means of each combination of spatial frequency, modulation, noise passband, and bandwidth/line rate were calculated and plotted. To describe quantitatively the results, simple linear regressions were calculated with modulation predicting threshold noise level at each spatial frequency and at each of the eleven bandwidth/line rate-noise passband combinations. Multiple linear regressions of spatial frequency and modulation predicting threshold noise level were also obtained at each bandwidth/line rate-noise passband combination.

Figures 4 through 7 show the relationship between square wave modulation and the threshold input noise level given in rms millivolts for the four noise passbands used in the 32 MHz bandwidth/1225 line system.

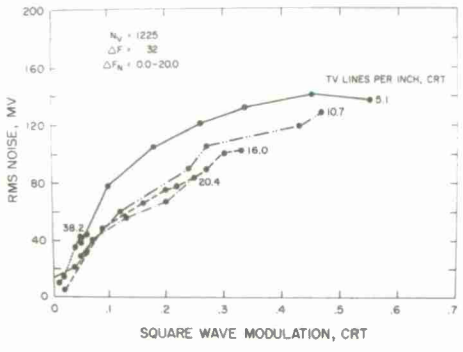



Figure 4. Detectability Threshold Means

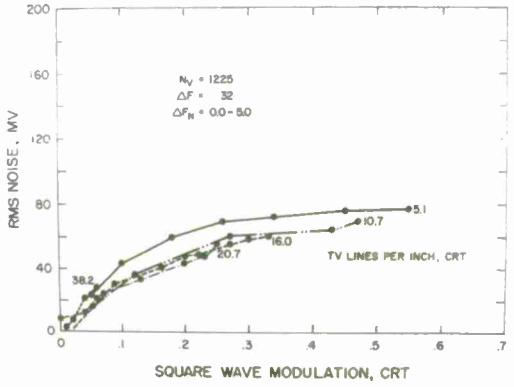



Figure 5. Detectability Threshold Means

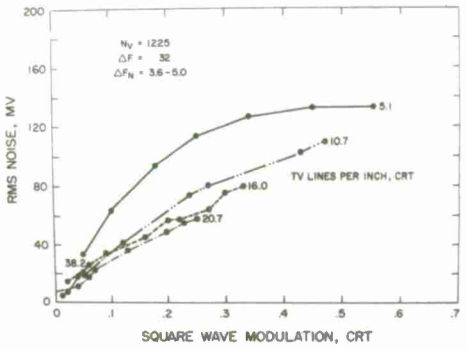



Figure 6. Detectability Threshold Means

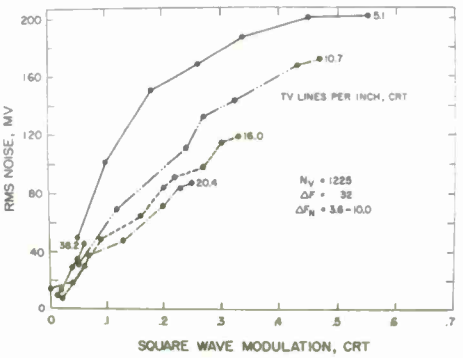



Figure 7. Detectability Threshold Means

Figur 4 is for the 0 to 20 MHz noise passband. The relationship among the parameters is as expected. At any spatial frequency, as modulation increases the threshold rms noise also increases. As spatial frequency increases (target size diminishes), the threshold noise declines. Note that for the largest target, spatial frequency equal to 5.1 TV lines/inch, the threshold curve is nonlinear. In the segment where modulation is less than

about 0.3, any increase in modulation is matched by an increase in the noise threshold. Beyond this segment, an increase in modulation allows little or no increase in noise. A plateau of sorts has been reached. For the largest targets with higher contrast, most of the subjects commented that the whole target disappeared at the point in increasing noise where they could no longer determine the existence of three separate bars. In other words, the target was totally obliterated at the noise threshold. In other targets, the target remained visible as a blur as noise was increased above the threshold even though three separate bars could not be seen.

Notice also in this plot, as well as those following, that the maximum modulation plotted at each spatial frequency declines with increasing spatial frequency. This is an illustration of the roll-off of the modulation transfer factor as spatial frequency increases.

Although targets of seven spatial frequencies were used at each of the eleven system combinations, only the lower spatial frequencies are shown in these graphs. Even with a fair range of modulation on the photographic prints of the higher spatial frequencies, the displayed modulation range is quite small due to the TV system MTF roll-off, yielding mean threshold curves of only two or three points. For clarity, these curves of limited usefulness were not plotted.

The detectability threshold curves for the 0 to 5 MHz noise passband given in Figure 5 show the same ordering of points and general shape as those in Figure 4. These threshold noise levels are only about half as great as those at the 0 to 20 MHz passband. Since the 0 to 5 MHz band is contained in both of these noise passbands, it must mean that the 0 to 5 MHz noise band contains the most detrimental noise frequencies. This is reasonable, since the lower frequencies will cause the larger "snow flakes." The energy expended in the 5 to 20 MHz noise frequencies apparently has much less effect on target detectability.

The detectability thresholds for the 3.6 to 5.0 MHz noise passband given in Figure 6 are slightly lower than those of the 0 to 20 MHz noise passband. If the noise in the 3.6 to 5.0 MHz region were an important contributor to the impairment of this sort of target detection then the rms noise voltages required to reach threshold conditions would be substantially less than that rms noise required at the 0 to 20 MHz noise passband. But they are not substantially less. More important, the threshold rms noise voltage for any target is much higher at 3.6 to 5.0 MHz noise passband than at the 0 to 5 MHz noise passband, indicating the important noise frequencies are between 0 and 3.6 MHz.

Because noise in the 3.6 to 5.0 MHz passband was relatively ineffectual in degrading targets, then noise energy in the 3.6 to 10 MHz is even more wasted. These thresholds, shown in Figure 7, are the highest in the 32 MHz bandwidth/1225 line group.

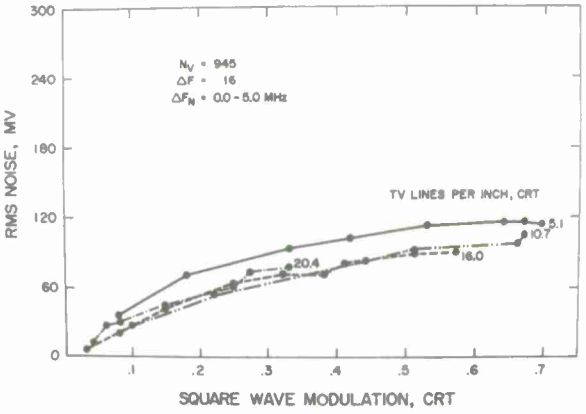



Figure 8. Detectability Threshold Means

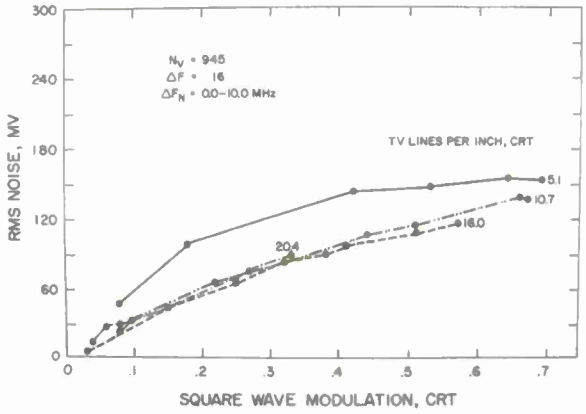



Figure 9. Detectability Threshold Means

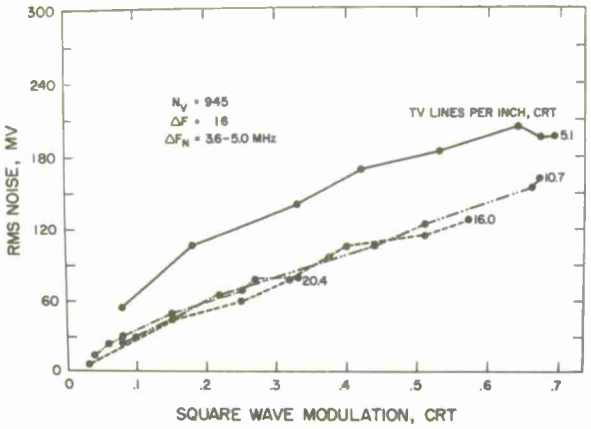



Figure 10. Detectability Threshold Means

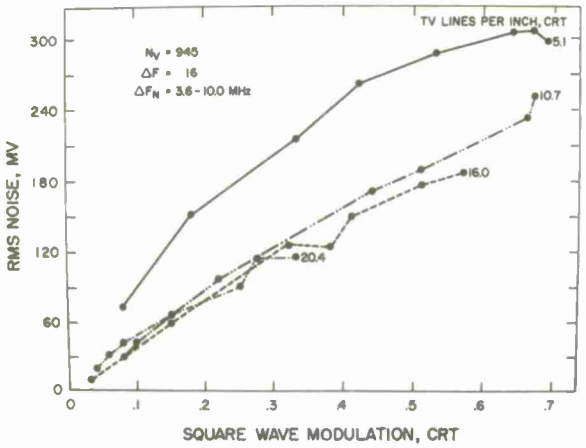



Figure 11. Detectability Threshold Means

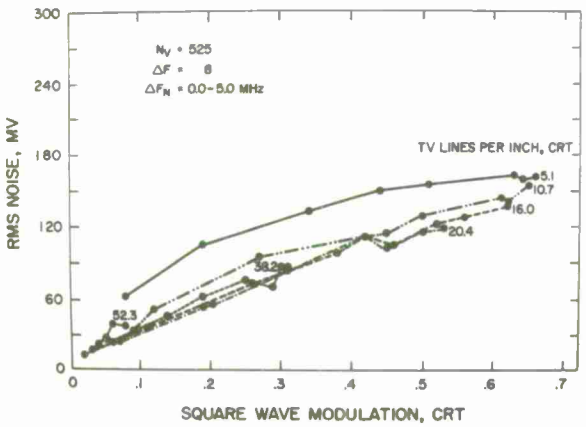



Figure 12. Detectability Threshold Means

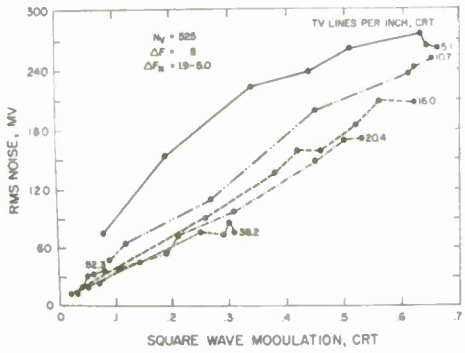



Figure 13. Detectability Threshold Means

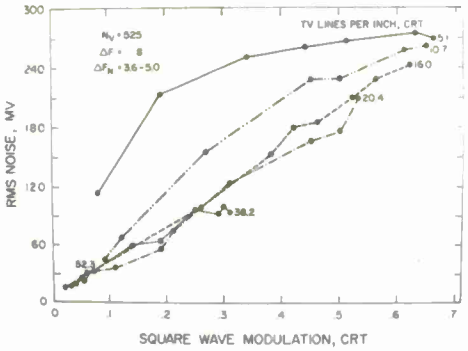



Figure 14. Detectability Threshold Means

The detectability thresholds for the 16 MHz/945 line and 8 MHz/525 line systems are presented in Figures 8 through 14. The ordering and shape of the curves is similar to those of the 32 MHz/1225 line systems. The detectability thresholds in the 8 MHz/525 line system with a 1.9 to 5.0 MHz noise passband are considerably higher than those for the 0 to 5 MHz noise passband at the same bandwidth/line rate. This result suggests that the most detrimental noise frequencies are below about 2 MHz.

At the 525 line rate, on a 10 in. x 14 in. monitor, 2 MHz converts to about 10 TV lines/inch at the monitor. As can be seen in Figure 1, this is the point where the square-wave response for the 8 MHz/525 line system starts to decline. One might conclude that the reduced contribution of noise frequencies greater than 2 MHz is due to the system response roll-off above 2 MHz; that is, frequencies greater than 2 MHz have less amplitude due to the monitor square-wave response.

This is only a partial explanation of the importance of low frequency noise. At the 1225 line rate, the 3.6 to 5.0 MHz noise passband falls between 7.5 and 10.0 TV lines/inch spatial frequency at the monitor. The square-wave response for the 32 MHz/1225 line system is relatively flat to about 10 TV lines/inch at the display although the square-wave modulation response is only about .67. So, the elevation of the noise thresholds for the 3.6 to 5.0 and 3.6 to 10.0 noise passbands is due to both the differential roll-off of the higher frequency noise by the system square-wave response and to the inherent detriment of lower frequency noise, i.e., noise less than about 2 MHz. The rms noise scale changes with each bandwidth/line rate combination.

## Discussion

Regression. Simple and multiple linear regressions were originally applied to the data for several reasons. First, a simple algebraic description of the results was sought to convert the data to a form useable in MTFA calculations. The original data were in the form of threshold noise level as a continuous variable dependent on discrete values of spatial frequency and modulation for each noise passband-bandwidth/line rate combination. For each combination, an applied linear regression is of the form

$$a(SF) + b(M) + c = \sigma_N \qquad (1)$$

where SF is the spatial frequency in TV lines/inch, M is modulation, $\sigma_N$ is the threshold noise level in rms millivolts, and a, b, and c are constants of the regression.

Also, the MTFA requires threshold curves relating spatial frequency and modulation with noise level at a discrete parameter. The required algebraic form is

$$a(SF) + c = M \qquad (2)$$

with $\sigma_N$ equal to some particular constant. Of course the regression equation (1) can be solved for modulation (M) to satisfy the MTFA requirement. Figure 15 plots, as an example, the multiple linear regression for the 32 MHz/1225 line, 0 to 20 MHz noise passband data.

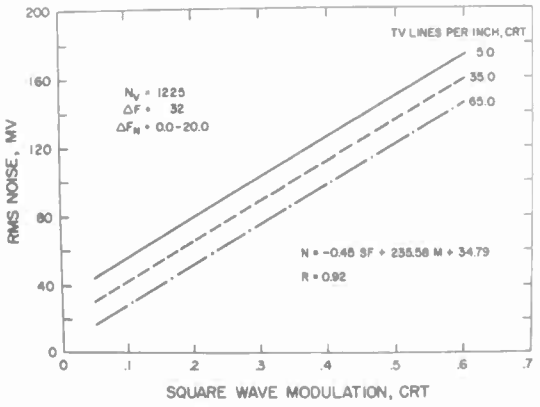



Figure 15. Detectability Threshold Multiple Linear Regression

A good question is the legitimacy of the linear, rather than nonlinear, regressions. Linear regressions were originally fit for speed and simplicity As shown in the summary of the multiple linear regressions, Table 1, the minimum correlation coefficient is .90; at least 81% (100 x $.90^2$) of the variance is predicted by the linear equation. Naturally, all F tests for regression effect proved highly significant. This result does not indicate that the linear model is the best-fitting one, however. An F test for lack of fit applied to the simple linear regression for 32 MHz/1225 line, 0 to 20 MHz noise passband, and spatial frequency equal to 5.1 TV lines/inch was significant ($p$ less than .001) indicating that the linear model was incorrect. In a rash attempt to find a better model, a stepwise multiple regression was applied to 27 different transformations of spatial frequency and modulation using the 32 MHz/1225 line, 0 to 20 MHz noise passband results. Using F = .01 for inclusion and F = .005 for deletion, the regression stopped after 12 steps at a multiple correlation coefficient of .94, indicating that no further increase of the coefficient was possible by either adding or deleting any of the 27 transformations. Since the multiple r for the linear model is .92 (Table 1), the improvement in percentage of the variance predicted by the regression is 3% [$(.94)^2$ = .88, $(.92)^2$ = .85]. Such a small improvement does not seem to warrant the use of an equation of twelve off-beat variables predicting threshold noise. This does suggest that 12% of the

data variance is truly random and unpredictable. On the assumption that the models underlying the other ten noise passband-bandwidth/line rate combinations are of similar form, the linear equations can be used with the realization that they are not perfect, but useful, very close approximations.

Grouping is especially evident in the 10.7, 16.0, and 20.4 TV line/inch threshold curves for the 16 MHz/945-line system. The equality of the simple linear regressions for these threshold curves should be treated further and, if significant, explained.

| $\Delta f/N_v$ | $\Delta f_n$ | R | Least Squares Regression Equation N = ____ SF + _____ M + _____ | | |
|---|---|---|---|---|---|
| 32 MHz/1225 | 0.0 - 20.0 MHz | .92 | -.45 | 235.58 | 34.79 |
| | 0.0 - 5.0 MHz | .94 | -.32 | 127.22 | 22.83 |
| | 3.6 - 5.0 MHz | .93 | -.32 | 225.79 | 20.08 |
| | 3.6 - 10.0 MHz | .93 | -.48 | 343.47 | 31.46 |
| 16 MHz/945 | 0.0 - 5.0 MHz | .95 | -.48 | 136.88 | 29.30 |
| | 0.0 - 10.0 MHz | .90 | -.54 | 184.95 | 33.29 |
| | 3.6 - 5.0 MHz | .94 | -.43 | 237.91 | 23.71 |
| | 3.6 - 10.0 MHz | .91 | -.58 | 370.03 | 29.98 |
| 8 MHz/525 | 0.0 - 5.0 MHz | .91 | -.56 | 190.03 | 40.99 |
| | 1.9 - 5.0 MHz | .91 | -.85 | 325.51 | 44.14 |
| | 3.6 - 5.0 MHz | .91 | -1.17 | 337.51 | 62.60 |

Table 1. Multiple Linear Regression

Spot Size Limitation. The order of square-wave response functions for these three systems (Figure 2) is reversed from expectations, due largely to the spot size limitation of the particular monitor used here. Although spot size measurements were not made, it is very likely that the spot size increased with the increased beam current used to achieve the faster writing speed of the higher line rates. This offers a possible explanation for the increase in threshold noise level with decreasing bandwidth/line rate. If, for the sake of argument, the spot sizes for the 8 MHz/525-line system and for the 32 MHz/1225-line system were equal, then the impairment due to a given amount of inserted noise would be greater in the 32 MHz/1225-line system since the greater writing speed results in a larger "snow flake" for any positive noise pulse. This explanation needs empirical evaluation.

Threshold Grouping. In many of the mean threshold figures, it appears that several of the threshold curves coincide.

## SYSTEM PHOTOMETRY

Besides describing quantitatively the target modulation and spatial frequency put into the system and the displayed modulation and spatial frequency at the monitor, the relationship between the electrical noise inserted into the system and the displayed luminous noise was also sought.

Input spatial frequency was physically measured on the photographic prints. The modulation of the targets on these prints was measured using a microphotometer with a 25 x 2500 micron scanning-slit eyepiece. Spatial frequencies of the targets at the display were calculated from the input spatial frequency and the system magnification. A sample of targets was measured at the display to verify these calculations. Target modulation at the display was found using the same microphotometer and scanning eyepiece as used in the input modulation measurement of the photographic prints. This displayed

modulation had to be measured at each different bandwidth/line rate combination since the MTF's changed.

Fortunately, the output rms voltmeter of the noise generator proved to be accurate as checked by a Ballantine Model 323-01 true rms voltmeter, and its readings were used for the input noise level. Corrections for attenuation of the noise passband were applied in the data analysis.

Obviously, the size of the scanning spot is also an important parameter of any line-scan display. An attempt to make this measurement was made using a high-efficiency microphotometer with a double-slit aperture (.003 x .400 inch with .150 inch spacing at the CRT). Theoretically, the spot passing the two slits should give two peaks on an oscilloscope displaying the output of the photomultiplier. From these two peaks, the spot size can be found. In this particular CRT, the persistence of the P-4 phosphor was so long relative to the speed of the spot that only one peak was obtained on the oscilloscope for each passing of the spot across the two slits.

This photomultiplier tube output was a good measure of the variation of the spot luminance due to inserted noise, however. The photograph of one of these oscilloscope images is given in Figure 16. The output of the photometer was taken directly from the dynode of the photomultiplier tube. Evidence to the absence of saturation of the photomultiplier tube is given by the lack of limiting seen in the peaks in Figure 16. With the photomultiplier tube operating in its linear region and the results displayed on a calibrated Tektronix 7403N oscilloscope, the peak height is linearly proportional to the luminance of the spot as it passes the aperture. The standard deviation of the peak heights when converted to footlamberts, is the rms luminance of the spot and can be compared with the rms noise voltage inserted.

The results are given in Figures 17 through 27. These graphs show, for each bandwidth/line rate-noise passband combination of the threshold experiment, the relationship between the inserted noise in rms volts and the displayed noise in arbitrary units of rms luminance. These measurements were made against both large white and black areas. The range of inserted noise at each system combination was determined by the approximate range of the detectability thresholds at these system combinations.

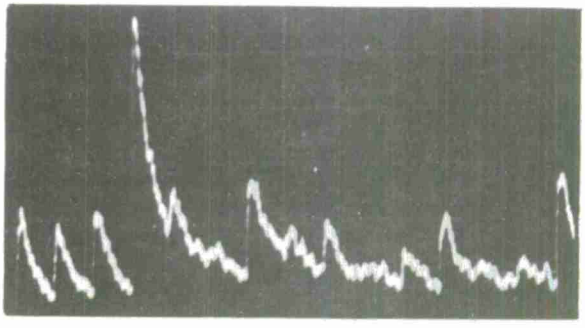


Figure 16. Oscilloscope Image of the Photomultiplier Output Measuring Displayed Noise

Before further discussion of Figures 17 through 27, it should be pointed out that there are data missing in Figures 21, 22, and 23. Some points near zero-inserted noise are not plotted for simplicity. Also, the abscissa scale is 0 to .25 rms volts in Figures 14 through 20 and 0 to .50 rms volts in Figures 21 through 27. Each point in the 32 MHz/1225-line system combinations is based on about 36 measured peaks, each point in the 16 MHz/945 line plots is based on about 28 peaks, and about 16 peaks were measured for each plotted point in the 8 MHz/525 line data.

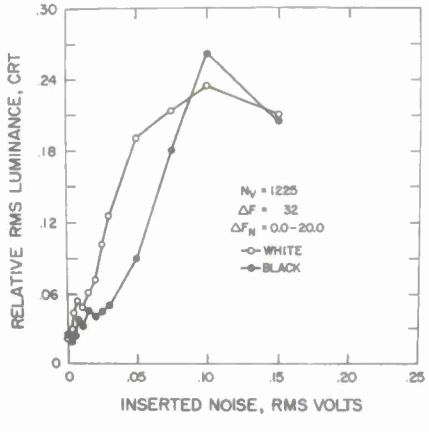



Figure 17. Rms Luminance as a Function of Inserted Noise

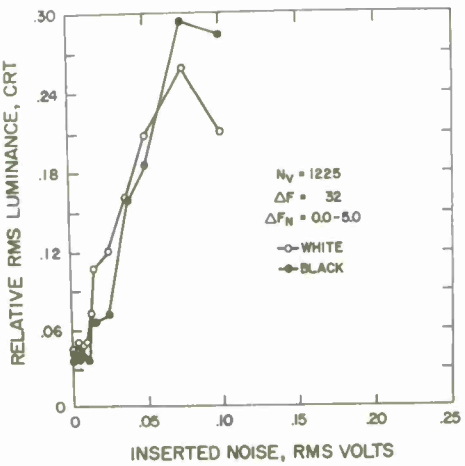



Figure 18. Rms Luminance as a Function of Inserted Noise

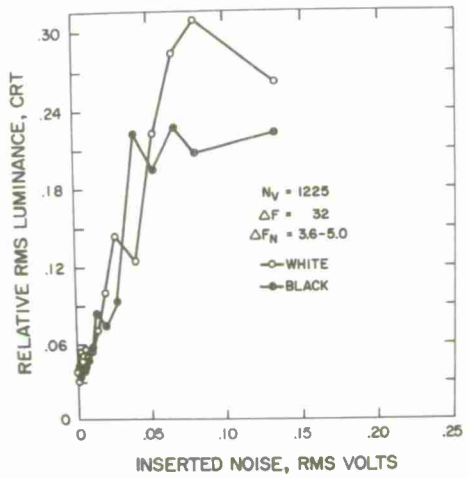



Figure 19. Rms Luminance as a Function of Inserted Noise

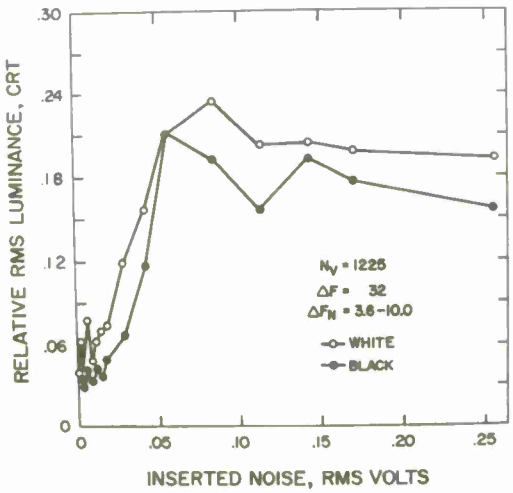



Figure 20. Rms Luminance as a Function of Inserted Noise

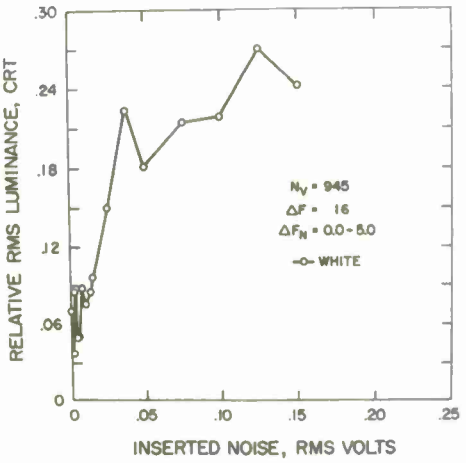



Figure 21. Rms Luminance as a Function of Inserted Noise

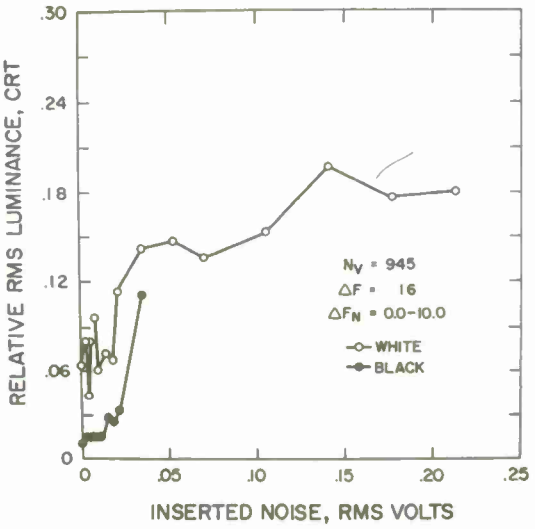



Figure 22. Rms Luminance as a Function of Inserted Noise

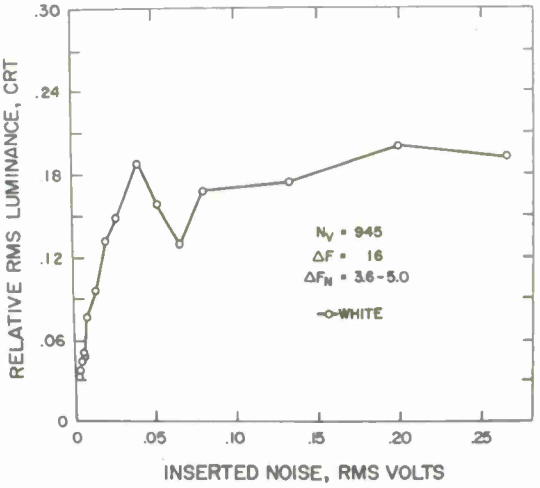



Figure 23. Rms Luminance as a Function of Inserted Noise

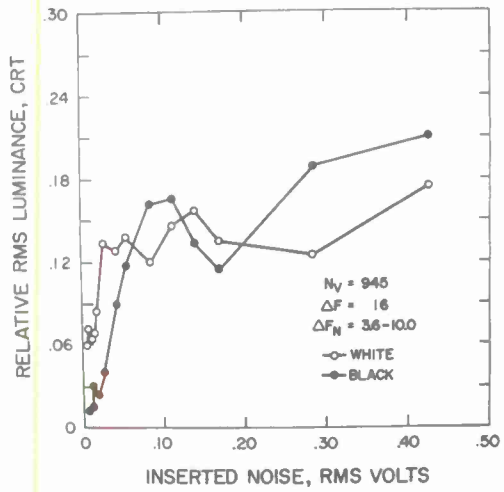



Figure 24. Rms Luminance as a Function of Inserted Noise

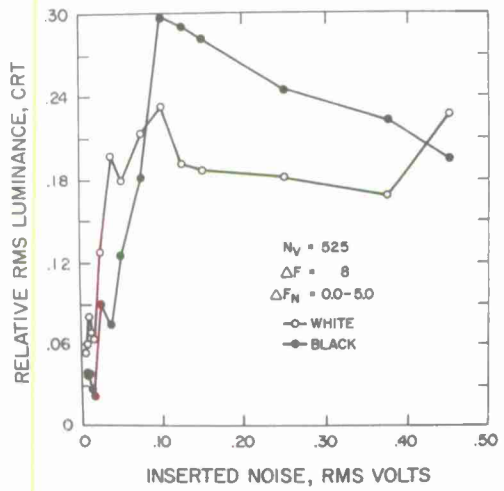



Figure 25. Rms Luminance as a Function of Inserted Noise

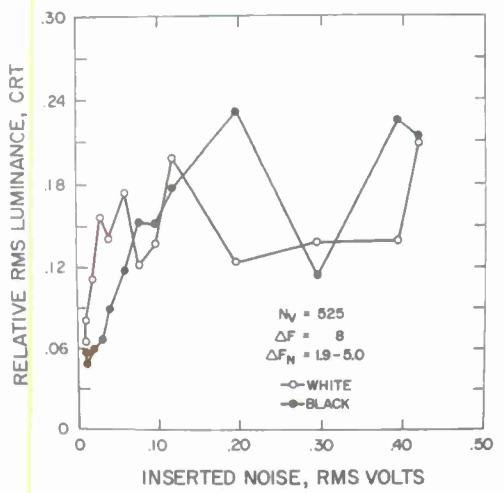



Figure 26. Rms Luminance as a Function of Inserted Noise

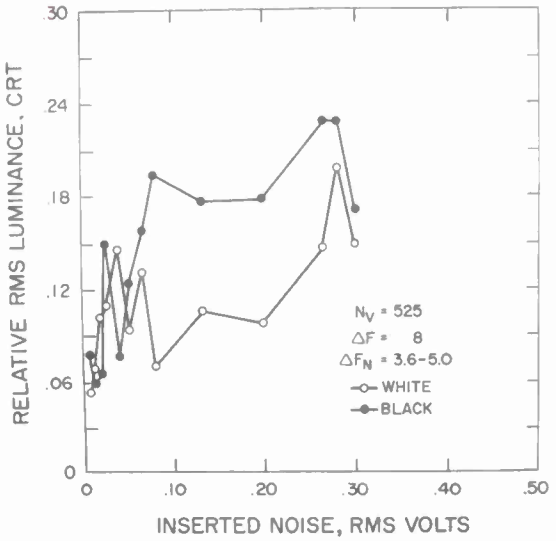



Figure 27. Rms Luminance as a Function of Inserted Noise

An inspection of these graphs indicates that application of any linear relation is inappropriate. In most cases, the curves are made up of two linear portions. For instance, in Figure 20 (32 MHz/1225 line, 3.6 to 10.0 MHz noise passband), increasing inserted noise up to about .05 rms volts results in a significant increase in the rms luminance. After this point, increasing inserted noise results in a slight decline in luminance variation.

Another consideration in this interpretation of the results is that the mean luminance, and thus the modulation, is not constant with increasing amounts of inserted noise. As the inserted noise increases, the mean luminance of a white area remains fairly stable but the mean luminance of a black area increases at the same rate as the rms luminance to about the point of peak rms luminance at which time both the rms luminance and the mean luminance stabilize.

The relationships among rms luminance, mean luminance, detectability thresholds, and target modulation need further study. There may also be an interaction here between the stability of the mean luminance and rms luminance at higher noise levels and the spot-size-limited nature of this monitor.

Because the purpose of this investigation into the display photometry was specification of display parameter values involved in finding the detectability thresholds, photometric noise investigation was limited and should certainly not be considered definitive.

## CONCLUSION

As expected, at any system and noise passband combination, increasing modulation at a particular spatial frequency allowed an increase in the noise threshold; and at any particular modulation, increasing spatial frequency brought about a decrease in the detectability threshold.

The ordering of the square-wave responde curves caused a corresponding ordering of the detectability thresholds. The lowest overall square-wave response was that of the 32 MHz/1225 line system and the lowest overall noise detectability thresholds were also for the 32 MHz/1225 line system.

Noise passband is very important in setting the detectability thresholds. The data from the different noise passbands indicate that lower frequency noise, less than 2 MHz, caused the greatest increase in the noise thresholds.

In the Introduction, it was stated that the principal objective of this work was generalization of the detectability thresholds to a variety of systems. The achievement of this objective is probably somewhat limited due largely to the spot-size-limited display and, in part, to the number of target spatial frequencies finally used. The degree of limitation from the spot-size problem should be investigated. It is apparent that the scanning spot size and shape are as important in specifying line-scan display image quality as are line rate and bandwidth. It is also interesting that the square-wave response curves gave proper consideration to this spot-size deficiency in ranking the three systems.

Finally, the photometry of displayed noise reported here was done only to quantify the stimuli used to find the detectability threshold. There is a definite further need to find the relationship between electrical noise and actual displayed noise. Certainly, it is true that research which assumes linearity between inserted rms electrical noise and displayed rms luminous energy noise may be in error, and should be reevaluated.

## ACKNOWLEDGMENTS

This research was supported by Contract Number F33617-71-C-1739 between VPI & SU and the Human Engineering Division, 6570th Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio.

## REFERENCES

Borough, H. C., Fallis, R. F., Warnock, T. H., and Britt, J. H. Quantitative determination of image quality. Boeing Report d2-114058-1, May 1967.

Charman, W. N. and Olin, A. Image quality criteria for aerial camera systems. Photographic Science and Engineering, 1965, 9, 385-397.

Coltman, J. W. and Anderson, A. E. Noise limitations to resolving power in electronic imaging. Proceedings of the IRE, 1960, 858-865.

Snyder, H. L. Detection and identification performance in air-to-ground visual search. Invited address to the Optical Society of America, Hollywood, Florida, October 1970.

Snyder, H. L. A unitary measure of video system image quality. Paper presented at the Target Acquisition Symposium, Orlando, Florida, November 1972.

# A UNITARY MEASURE OF VIDEO SYSTEM IMAGE QUALITY

by

Harry L. Snyder, PhD
Professor and Director, Human Factors Laboratory
Virginia Polytechnic Institute and State University
Blacksburg, Virginia 24061

OFFICE OF NAVAL RESEARCH
TARGET ACQUISITION SYMPOSIUM
NAVAL TRAINING CENTER, ORLANDO, FLORIDA / 14,15,16 NOVEMBER 1972

# A UNITARY MEASURE OF VIDEO SYSTEM IMAGE QUALITY

Harry L. Snyder, Ph.D.

Professor and Director, Human Factors Laboratory
Virginia Polytechnic Institute and State University
Blacksburg, Virginia 24061

## ABSTRACT

A single measure of image quality, the square-wave modification of the modulation transfer function area, is shown to correlate highly with observer performance in extracting information from both static and dynamic imagery. This paper describes this unitary measure of image quality and discusses the results of the experiments which establish its validity.

## INTRODUCTION

During the past two decades, over 300 laboratory, analytical, and field studies have been performed to assess the relationship between variation in line-scan display image parameters and some measure of observer performance. Conclusions drawn from critical reviews of these studies (e.g., Snyder, et al., 1967; Hairfield, 1970; Filleau, 1970) have indicated that cross-study comparisons are virtually impossible. Variation in specific system design parameters, or in the manner by which display image quality is synthetically manipulated, is often incompletely controlled, so that concomitant variation in the several contributing sources of image quality results.

Many experimental variables have been shown to have a significant effect upon operator information extraction performance. Among these are video bandwidth, number of scan lines, field of view, signal-to-noise level in the video, frame integration time, display contrast, scene rate of movement, and display size. For the most part, individual experiments have tended to examine the effects of one, two, and sometimes three of these variables. However, due to the inherent interaction (nonindependence) among these variables, both in terms of engineering design and human performance, quantitative combination of the results is hazardous even in the presence of good experimental control and measurement. In the absence of such control, any a posteriori attempt to combine the results is merely foolish.

Because of these gross conflicts and inconsistencies in the experimental literature dealing with the effects of individual system parameters, recent efforts have been oriented toward the development of (1) analytical expressions of overall image quality, such as those discussed by others here at this meeting and (2) experimental evaluations of logically-derived summary measures of image quality. The remainder of this paper will discuss the present approach and limitations of data pertaining to one such summary measure of image quality.

Any summary measure of image quality, to be useful, must be (1) easily measured for existing imaging systems, (2) quantitatively predictable, analytically, for future imaging systems at the paper design stage, and, in particular, (3) highly correlated with (or validated by) empirically-determined operator performance under the dynamic operational conditions of interest for the specified mission. To date, one summary measure of image

quality which shows promise for meeting these criteria is the Modulation Transfer Function Area.

Modulation Transfer Function Area (MTFA)

Originally proposed by Charman and Olin (1965) who termed it the threshold quality factor, and renamed by Borough, Fallis, Warnock, and Britt (1967), the MTFA concept has been evaluated in two experimental photointerpretation situations and demonstrated to relate strongly to the ability of image interpreters to obtain critical information from reconnaissance photographic imagery. In its original form, the MTFA was proposed as a unitary measure of photographic image quality which contains "the cumulative effect of the various stages of the atmosphere-camera-emulsion-development-observation process, the 'noise' introduced in the perceived image by photographic grain, and the limitations imposed by the physiological and psychological systems of the observer" (Charman and Olin, 1965, p. 385). While this measure was originally developed for direct photographic systems, its generalization to electro-optical systems is possible with some modifications.

The MTFA is derived in such a manner as to make use of the modulation transfer function (MTF) of the imaging system, thereby retaining the analytical convenience of component analysis based upon sine-wave response characteristics. In addition, it attempts to take into account other variables critical to the imaging and interpreting problem, such as exposure, the characteristic curve, granularity, the human observer capabilities and limitations, and the nature of the interpretation task. For the electro-optical system, the first three of these variables can be considered analogous to detector irradiance level, gamma (typically unity), and noise, respectively.

Figure 1 shows that the MTFA is the area bounded by the imaging system MTF curve and the detection threshold curve of the total system, including the eye. The MTF curve for the imaging system is obtained in the conventional manner, while the detection threshold curve requires several assumptions regarding the human operator. Specifically, it is assumed that the viewing conditions are optimum, and that threshold detection of any target in the imaged display is a function of the target image contrast modulation, the noise in the observer visual system, and the noise in the imaging system exclusive of the observer. It should be noted that the crossover of the two curves in Figure 1 represents the limiting resolution of the system for a sine-wave target.

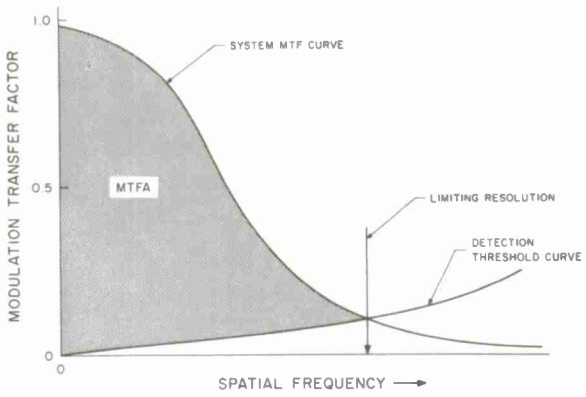



Figure 1. Modulation transfer function area (MTFA).

At low spatial frequencies, the threshold detection curve is dependent upon the properties of the human visual system, as shown in Figure 2. At higher spatial frequencies, the effect of imaging system noise becomes important. For the photographic case, this imaging system noise is equivalent to granularity. It is assumed further that the eye's contrast threshold is 0.04, so that this target image contrast modulation must be realized at the display for the target to be detected, regardless of the contrast modulation of the target object.

Figure 2 illustrates the normalized detection threshold curve, which must be adjusted both vertically and horizontally for a specific set of conditions. The curve is repositioned vertically, by moving it upward for decreases in scene object modulation, and by moving it downward for increases in system gamma. The detection threshold curve is also moved horizontally to the left with increases in film granularity, or noise.

Details of these manipulations are given by Borough, et al. (1967) and Snyder (in press).

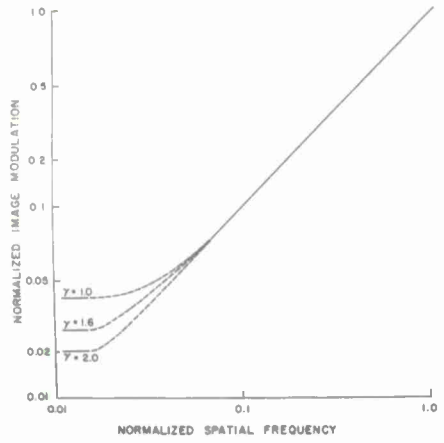



Figure 2. Generalized detection threshold.

Following positioning of the threshold detection curve, the MTFA is calculated by:

$$MTFA = \int_0^{N_1} \left(R_N - \frac{M_t(N)}{M_o(N)}\right) dN \qquad (1)$$

where $N_1$ = the spatial frequency at which the MTF curve crosses the detection threshold curve (limiting resolution)

$R_N$ = the MTF value at spatial frequency N

$M_t(N)$ = the normalized threshold value, as shown on Figure 2

$M_o(N)$ = the inherent object contrast modulation

It should be noted that the detection threshold curve, as described here, is akin to such concepts as contrast sensitivity (Campbell and Green, 1965), sine-wave response (DePalma and Lowry, 1962; Lowry and DePalma, 1961; Bryngdahl, 1966), and demand modulation function.

## Evaluation of the MTFA for Photographic Imagery

To date, two empirical evaluations of the MTFA concept have been conducted, both using photographic imagery. In the first study, an attempt was made to relate the MTFA to subjective estimates of image quality obtained from a large number of trained image interpreters. In the second of these experiments, actual information-extraction performance data were obtained, as well as subjective estimates of image quality, and both measures were compared with the MTFA values of the imagery.

In the Borough, et al. (1967) experiment to evaluate MTFA, the purpose was to determine whether a strong relationship existed between MTFA and subjective image quality. This limited evaluation was imposed simply to reduce data collection costs in the event that the MTFA measure proved fruitless. Nine photographic reconnaissance negatives were used as the basis for laboratory-controlled manipulation of image quality. Each of the scenes was printed in 32 different MTFA variants, determined by four different MTF's, three levels of granularity, and three levels of contrast, as illustrated in Figure 3. Four cells of the matrix were deleted because their MTFA values corresponded to others in the 32-cell matrix. The MTF curves are illustrated in Figure 4.

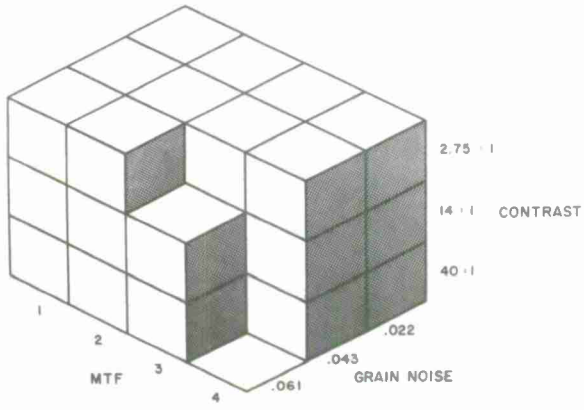



Figure 3. Production of MTFA conditions.

The resulting 288 transparencies (9 scenes by 32 variants/scene) were used in a partial paired-comparison evaluation by 36 experienced photointerpreters. The subjects were asked to select the photo of each pair that had the best quality for extraction of intelligence

information. All pairs were composed of two variants of the same scene; each subject made a total of 256 comparisons, for a grand total for all subjects of 36 X 256 = 9216 judgements.

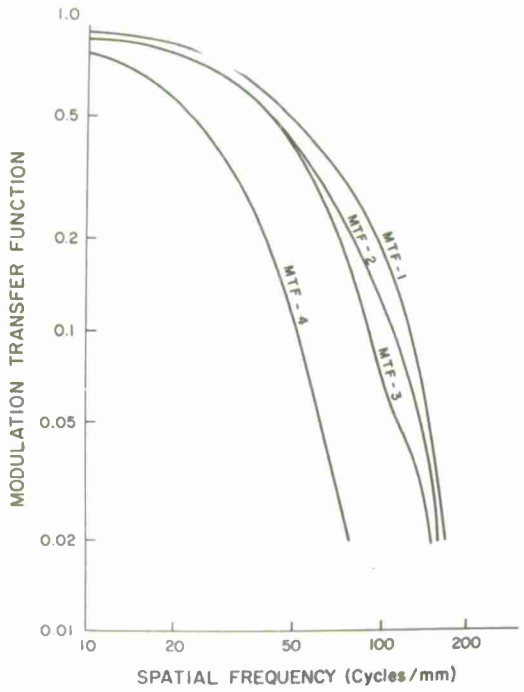



Figure 4. Average MTF curves produced by edge-response method.

Correlations were obtained between the subjective image quality rating (derived from the paired-comparisons) for each of the 32 variants and several physical measures of image quality. Most important to this discussion is the obtained mean correlation of 0.92 between MTFA and subjective image quality, which indicates that the MTFA is strongly related to subjective estimates of image quality.

The next experiment, by Klingberg, Elworth, and Filleau (1970), examined the relationship between objectively-measured information-extraction performance and the MTFA values. As a check on the results of Borough, et al. (1967), Klingberg, et al. (1970) also obtained subjective estimates of image quality, so that all three intercorrelations were evaluated.

The imagery used for this experiment was the same as that used by Borough, et al. (1967). A group of 384 trained military photointerpreters served as subjects. Each subject was given one variant of each of the nine scenes and asked to (1) rank the image on a nine-point interpretability scale, using utility of image quality for information extraction as the criterion, and (2) answer each of eight multiple-choice questions dealing with the content of the scene. The interpretability scale values were used to develop a subjective image quality measure for the 288 images, while scores on the multiple-choice interpretation questions were used to measure information extraction performance.

Figure 5 shows the scattergram between information extraction performance and MTFA for the 32 MTFA values. The resulting product-moment correlation, averaged across the nine scenes, is -0.93. (The minus value is due to the use of number of errors as a measure, which is inversely related to MTFA.) Further, the obtained product-moment correlation of 0.96 between MTFA and subjective quality (rank) agrees quite well with the correlation previously obtained by Borough, et al. (1967), of 0.92. The correlation between subjective quality and the number of errors is 0.97.

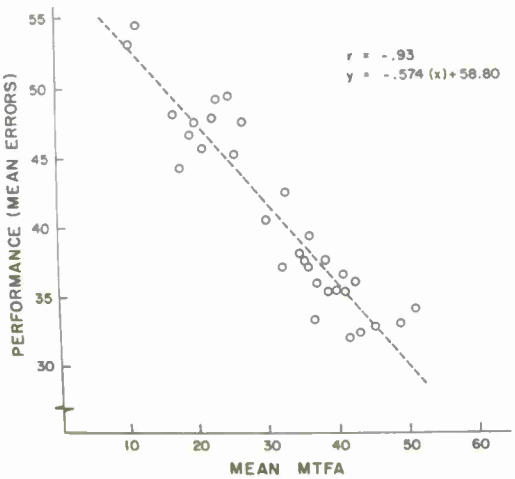



Figure 5. Scattergram, information extraction performance vs. MTFA.

These data show unequivocally that a measure of image quality based upon the excess of MTF over the threshold detection

level correlates highly with the ability of observers to obtain critical operational information from photographic imagery. The remaining question, of course, is whether the MTFA concept can be validly used to predict typical observer performance from an electro-optical line-scan display.

## THE MTFA AND LINE-SCAN DISPLAYS

The application of the MTFA concept to line-scan sensors and displays is not straightforward, either analytically or experimentally. First, there exists no threshold detection curve for the line-scan display situation in which the targets are sinusoidally varying in intensity. Second, there is no easy and perhaps valid derivation of the effects of variables such as the number of active scanning lines, the noise passband, the video passband, and the target modulation upon the detectability threshold curve. Generalization from the concepts offered by Charman and Olin (1965) cannot be made with any confidence. Finally, the Charman and Olin threshold curves imply optimum viewing by the subject, including optimum display size (image magnification), unlimited viewing time, ideal non-glare ambient illumination, and no annoying environmental conditions (e.g., noise, vibration, etc.).

For these reasons, our approach to employing the MTFA concept has been to modify it as needed to better fit the line-scan display case, both conceptually and from a measurement standpoint, and to determine empirically the appropriate threshold detection curves for our particular viewing conditions. R. L. Keesee's paper, at this Symposium, will discuss the methodology and results of our experiments performed to generate the threshold curves, while the remainder of this paper will present the concepts and experiments used to validate those concepts.

### Modification of the MTFA Concept

The use of a sine-wave pattern to generate both the MTF curve for a given system and to determine detection threshold curves is quite difficult, in part because it is difficult to obtain or produce accurate sinusoidally-varying intensity patterns on hard copy. For this reason, and also because it is reasonably easy to transform square-wave response data into approximately equivalent sine-wave response data, we have been using square-wave inputs exclusively. Our experimental technique is summarized in Figure 6. Simply, a variable parameter television system is used to view either 8x10-inch hard copy prints or a dynamically-varying scene contained on 35mm motion-picture imagery which was taken from the North American Aviation/Columbus terrain model at varying speeds, altitudes, fields of view, and depression angles (Snyder, 1967). System parameters such as video bandwidth, active line number, contrast, etc., can be selectively varied by controls at either the camera or the camera control unit. The noncomposite video is mixed with a shaped noise signal by a wide-bandpass video mixer, added to the sync and blanking signals, and sent to a variable-parameter 17-inch display located 40 inches from the subject's nominal eye position.

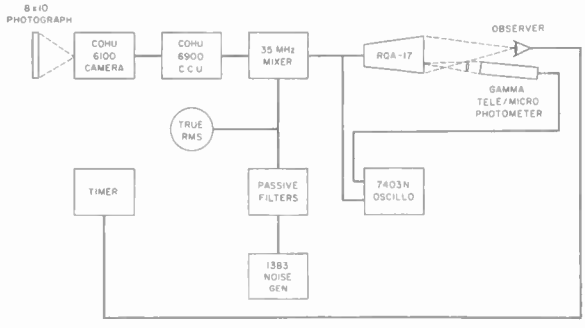



Figure 6. Simulation system block diagram.

The video signal is monitored at several places and displayed on a wide-bandwidth oscilloscope, the face of which is photographed for subsequent evaluation. These test points permit measurement of rms noise, peak-to-peak signal, signal plus noise waveform, and camera raw video output.

All camera and camera control unit adjustments are checked every 30 minutes during an experimental run to assure

stability both electrically and photometrically. The photometric quality of the display is measured with a telephotometer to assure constant gray-scale rendition, using an 11-level gray-scale chart. The photometric output of the system, relative to a square-wave 1951 USAF tribar target input, is measured photometrically with a scanning eyepiece by a microphotometer, recording the analog output on an X-Y plotter. Thus, the system square-wave response is measured photometrically, rather than electrically. As Keesee's paper shows, it is not necessarily safe to assume linearity between electrical input and photometric output of a cathode-ray tube. Our emphasis on the photometric output is simply because the visual stimulus is photometric, not electrical, and we are ultimately interested in measuring human performance.

Employing these techniques, and varying several system parameters, we can generate a variety of system square-wave response, $R_{SQ}(N)$, curves, the experimental results from which are discussed in the remainder of this paper.

## Evaluation of the Modified MTFA for Line-Scan Displays

In one experiment, five different noise levels were used with a 1225-line, 32-MHz bandwidth video system configuration. The system square-wave response $R_{SQ}(N)$ and the square-wave detection threshold curves are shown in Figure 7. The imagery displayed was an air-to-ground dynamic search scene, simulating an altitude of 10,000 feet, a groundspeed of 500 feet per second, a camera depression angle of 28 degrees, and a camera field of view of 30 degrees diagonal on a 3:4 (horizontal: vertical) aspect ratio. The video signal-to-noise levels, for a 100% contrast target, were 32, 9.14, 4.27, 2.91, and 2.13 decibels.

Twenty-eight tactical and strategic ground targets were selected, and 11 different subjects were used for each of the five noise levels, each subject searching for each of the 28 targets. The subjects' performance was measured in terms of the likelihood of a correct response (number of correct responses divided by 28) and the calculated range to the target at the time of a correct recognition response. All targets were pre-briefed, with the subjects having studied pictures of the targets prior to the experiment.

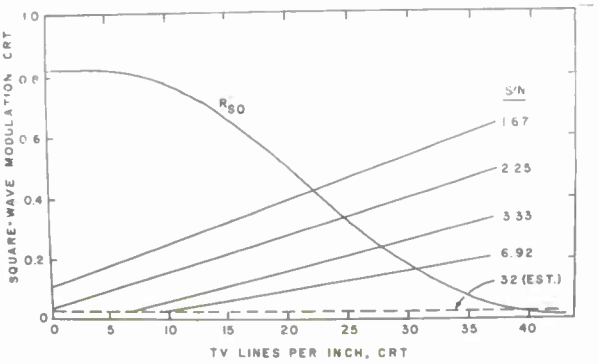



Figure 7. System $R_{SQ}(N)$ and threshold detection curves for five noise levels.

The results of the experiment, summarized, were that the modified MTFA measure ($MTFA_{SQ}$) correlates 0.965 with the likelihood of recognition, and 0.76 with the range of recognition.

In a second experiment, using static pictures of human faces, similar results were found. In this study, we were interested in the application of image quality to the recognition of a single face, as in a police-surveillance task from low-light-level video tape imagery. Static facial photos were placed before the television camera, and the subjects were required to identify the face on the display from a "rogues' gallery" set of 35 faces. Three different system configurations were used, with each configuration studied at five different noise levels, as illustrated in Figures 8 through 10. Table 1 summarizes the experimental design and conditions, indicating that each of five subjects was asked to recognize each of the 35 faces, in a random order, three different times, once for each system (under a different noise level). Order of presentation and selection of the facial stimuli for each subject and for each condition were appropriately counterbalanced.

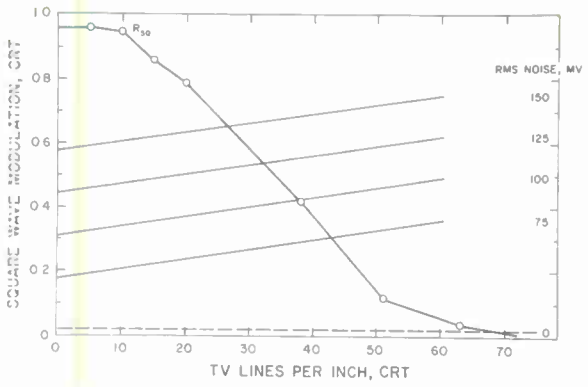



Figure 8. System $R_{SQ}(N)$ and threshold detection curves, 8/525 system.

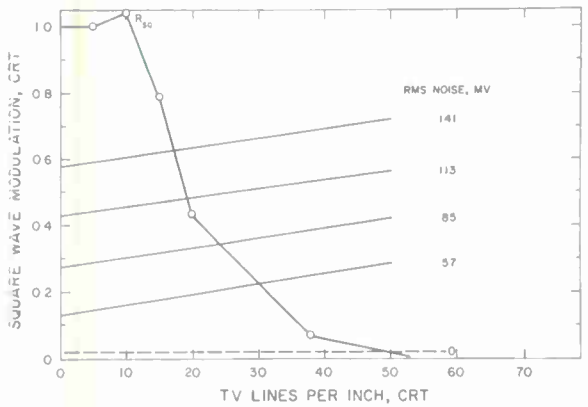



Figure 9. System $R_{SQ}(N)$ and threshold detection curves, 16/945 system.

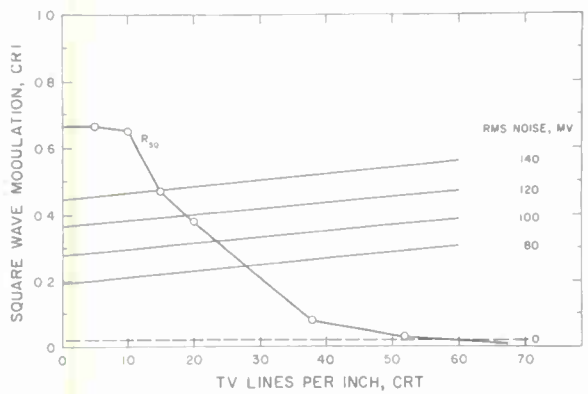



Figure 10. System $R_{SQ}(N)$ and threshold detection curves, 32/1225 system.

| | RMS Noise, Mv. | Subjects | | | | | $MTFA_{SQ}$ |
|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | |
| 8/525 | 0 | Subset 1 | Subset 2 | Subset 3 | Subset 4 | Subset 5 | 30.46 |
| | 75 | " 2 | " 3 | " 4 | " 5 | " 1 | 18.46 |
| | 100 | " 3 | " 4 | " 5 | " 1 | " 2 | 13.22 |
| | 125 | " 4 | " 5 | " 1 | " 2 | " 3 | 8.72 |
| | 150 | " 5 | " 1 | " 2 | " 3 | " 4 | 4.48 |
| 16/945 | 0 | " 6 | " 7 | " 8 | " 9 | " 10 | 21.09 |
| | 57 | " 7 | Etc. | Etc. | Etc. | Etc. | 15.23 |
| | 85 | " 8 | | | | | 11.47 |
| | 113 | " 9 | | | | | 8.33 |
| | 141 | " 10 | | | | | 5.64 |
| 32/1225 | 0 | " 11 | " 12 | " 13 | " 14 | " 15 | 14.65 |
| | 80 | " 12 | Etc. | Etc. | Etc. | Etc. | 7.37 |
| | 100 | Etc. | | | | | 5.33 |
| | 120 | | | | | | 3.67 |
| | 140 | | | | | | 2.36 |

Table 1. Experimental design, facial recognition experiment.

The results of this facial recognition experiment are shown in Figures 11 and 12. The product-moment correlation between the fifteen experimental conditions' $MTFA_{SQ}$ values and the probability of recognizing the face is 0.69, although the relationship is obviously nonlinear (Figure 11). Similarly, Figure 12 shows that the relationship between $MTFA_{SQ}$ and response time is orderly, but nonlinear, with a correlation of -0.67. Applying transformations to each plot to approximate linearity, and replotting the data as Figures 13 and 14, it is shown that the correlation between $\log_{10}MTFA_{SQ}$ and the probability of correct recognition is 0.87 (Figure 13) and the correlation between $\log_{10}MTFA_{SQ}$ and $\log_{10}$ response time is -0.92 (Figure 14). Both correlations are statistically significant at the 0.001 level of confidence. The nonlinear best-fit equation is of no consequence in terms of systematic prediction; it is only important that a high correlation exist to validate the concept of the metric.

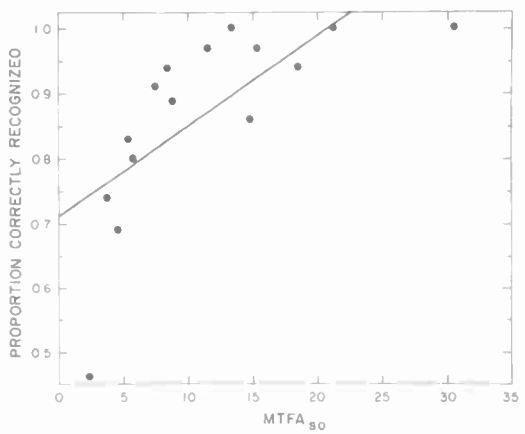



Figure 11. $MTFA_{SQ}$ vs. probability of correct recognition of faces.

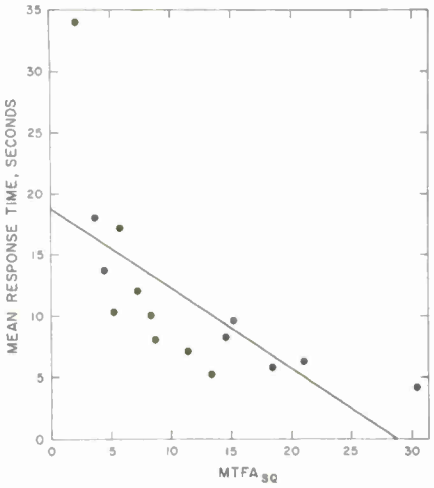



Figure 12. $MTFA_{SQ}$ vs. response time for facial recognition.

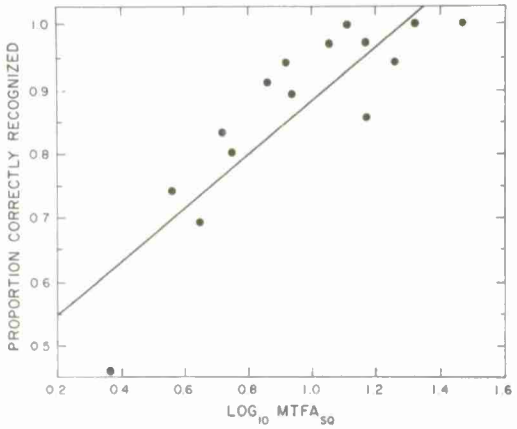



Figure 13. $MTFA_{SQ}$ vs. $\log_{10}$ probability of correct recognition.

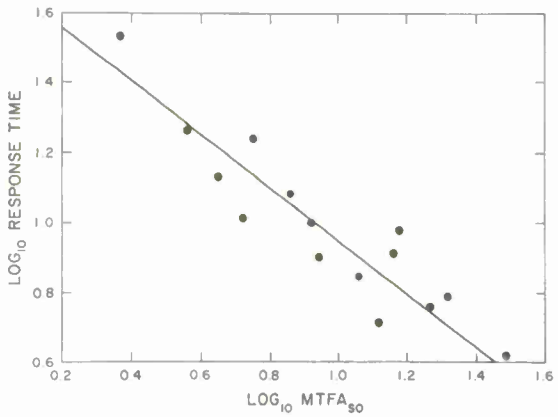



Figure 14. $Log_{10}$ $MTFA_{SQ}$ vs. $\log_{10}$ response time.

## CONCLUSION

In general, then, we can conclude that the revised metric based upon a square-wave response, $MTFA_{SQ}$, is a strong predictor of image quality for raster-scan displays, just as it is for continuous-sampled photographic imagery. In fact, under certain conditions, such as when system gamma is unity and the $MTFA_{SQ}$ is transformed into an equivalent sine-wave metric, the MTFA concept is very similar to the integral under the $SNR_D$ curve which was discussed by Rosell at this Symposium. Using a periodic target, such as a tribar, the two concepts permit prediction of the human observer's performance with high validity. Rosell's studies of threshold detection for static images, and our studies of recognition for both static and dynamic images both show that this type of metric, based upon the system square-wave or sine-wave response, relative to the observer's requirement for contrast at a given spatial frequency, are accurate predictors of observer performance. In general, whatever system gives the greatest excess of contrast (or modulation) over the eye's requirement at the spatial frequency of interest will produce the best performance.

## REFERENCES

Borough, H. C., Fallis, R. F., Warnock, T. H., and Britt, J. H., Quantitative determination of image quality. Boeing Company Report D2-114058-1, May 1967.

Bryngdahl, C., Characteristics of the visual system: psychophysical measurements of the response to spatial sine-wave stimuli in the photopic region. Journal Optical Society of America, 1966, 56, 811-821.

Campbell, F. W., and Green, D. G., Optical and retinal factors affecting visual resolution. Journal Physiology, 1965, 181, 576-593.

Charman, W. N., and Olin, A., Image quality criteria for aerial camera systems. Photog. Sci. Engin., 1965, 9, 385-397.

DePalma, J. J., and Lowry, E. M., Sine-wave response of the visual system. II. Sine-wave and square-wave contrast sensitivity. Journal Optical Society of America, 1962, 52, 328-335.

Filleau, C. R., Night and all-weather target acquisition: State-of-the-art review. Part II: Infrared and laser systems. Boeing Company Report D162-10116-2, April 1970.

Hairfield, H. W., Night and all-weather target acquisition: State-of-the-art review. Part III: Television and low-light-level television systems. Boeing Company Report D162-10116-3, May 1970.

Klingberg, C. L., Elworth, C. L., and Filleau, C. R., Image quality and detection performance of military interpreters. Final report by The Boeing Company under USAFOSR Contract F 44620-69-C-0128, April 1970.

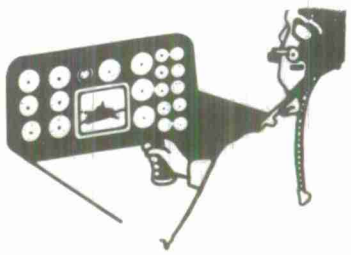
Lowry, E. M., and DePalma, J. J., Sine-wave response of the visual system. I. The Mach phenomenon. Journal Optical Society of America, 1961, 51, 740-746.

Snyder, H. L., Image quality and observer performance. In L. M. Biberman (Ed.) Perception of Displayed Information, Plenum Press, in press.

Snyder, H. L., et al., Low light level TV viewfinder simulation program. Phase A: State-of-the-art review. USAF Report AFAL-TR-67-293, November 1967 (S).

ACKNOWLEDGMENT

This research was supported by Contract Number F33617-71-C-1739 between VPI & SU and the Human Engineering Division, 6570th Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio.



# TARGET ACQUISITION THROUGH VISUAL RECOGNITION: AN EARLY MODEL

by

H. H. Bailey
The Rand Corporation
Santa Monica, California

OFFICE OF NAVAL RESEARCH
TARGET ACQUISITION SYMPOSIUM
NAVAL TRAINING CENTER, ORLANDO, FLORIDA / 14,15,16 NOVEMBER 1972

# TARGET ACQUISITION THROUGH VISUAL RECOGNITION: AN EARLY MODEL

H. H. Bailey
The Rand Corporation
Santa Monica, California

## ABSTRACT

One of the early attempts to provide a quantitative model for predicting the capabilities of an observer in looking for pre-briefed targets is described. This model is structured according to three distinguishable psychophysical processes: deliberate search over a fairly well-defined area; detection of contrasts (a subconscious retino-neural process); and recognition of shapes outlined by the contrast contours (a conscious decision based on comparison with memory). The probability that the three steps are completed successfully gives the probability of target detection. The foregoing is essentially a static model. Modifications to allow for the dynamics of flight approaching a target, and for individual target motion relative to the background, are also described.

## INTRODUCTION

Human vision is achieved by a remarkable electro-optical system that has evolved naturally to a form that is far more versatile than anything we have been able to build ourselves. Everyone is aware of the two classes of primary receptors, the efficient distribution of the high-resolution components, the huge range of luminance levels over which the system can function, and so forth.

There is an additional feature of the dynamic range capability that has been brought to light in a little-known fundamental paper (Ory, 1969). Using modern statistical decision theory, Ory was able to obtain an excellent fit to the 6-sec/8-position Tiffany data by making the following assumptions: primarily, the decision criterion is assumed to be adaptable, or programmed, and to depend on both the luminance (total number of neural excitations per unit time) and its fluctuations (square root of number of excitations per unit time); in addition, both the background and the target luminance and fluctuations are to be included, and the image extension due to both optical aberrations and the retinal neural organization must be accounted for. With these assumptions, all of the experimental data on the threshold luminance ($\Delta L\theta^2 t$) can be plotted on two curves, one for foveal-photopic vision and one for extrafoveal-scotopic vision. Each curve is a simple quadratic in the total fluctuation rate, and one of them has a small constant curve.

The interesting point here is that these curves define two regions, depending on whether the fluctuation or the luminance term is dominant; and these correspond to the DeVries-Rose and the Weber-Fechner regimes of the literature. These regimes are interpreted as follows: At low light levels, discriminable increments are limited by the noise level, $\Delta L \sim \sqrt{L}$, and system performance is determined by the signal-to-noise ratio, $\Delta L/\sqrt{L}$. As the luminance is increased, the number of discriminable levels rises as the square root, until another limit sets in--the number, n, of levels that our visual equipment can process simultaneously. From this point on, the discriminable increments are just the fraction 1/n of the maximum luminance, $\Delta L \sim L$, and system performance is limited by the contrast, $\Delta L/L$.

Now, with modern man-made electro-optical systems, despite the fact that the display is bright enough for the eye to

operate photopically, there is so much noise in any practical system that the overall performance is again noise-limited. The number of distinguishable grey levels on some of the old radar scopes was 3 or 4 at most, and now we struggle to get 7 or 8; but these are still well below the number (probably around 20) where the human processing system imposes its contrast limitation.

The conclusion is that signal-to-noise is the appropriate parameter for analyzing the capabilities of an observer viewing an electro-optical display; but it is still appropriate to use contrast when you look out the window.

Now I will go back and describe my old observer model. This was primarily intended for, and is best applied to, the contrast regimes. A description of this model was first published in 1970 (Bailey, 1970); however, it was developed during the mid-sixties and was used internally at Rand prior to its publication. This paper gives an abridged description of the model, emphasizing the concepts involved, together with two extensions which have been added since that time.

## GENERAL FORM OF THE MODEL

The performance of a human observer is often a very complicated function of many interacting variables. In order to simplify this difficult situation and yet stay close to reality, we consider explicitly the task of finding known and fixed objects in a complex field in a short time. This process, even when so restricted, is still complex, but it can be considered to consist of the following three distinct steps: deliberate search over a fairly well-defined area; detection of contrasts (a subconscious retino-neural process); and recognition of shapes outlined by the contrast contours (a conscious decision based on comparison with memory).

On the basis of assorted experimental data, three formulas can be devised for the probabilities of completing each of the three steps separately. It is postulated that the overall target recognition probability can be expressed by the product of these three terms. Accordingly, we establish the following definitions:

- $P_1$ is the probability that an observer, searching an area that is known to contain a target, looks for a specified glimpse time (viz., 1/3 sec) in the direction of the target with his foveal vision. $P_1$ is a function of the ratio of an acceptable search rate to that demanded of a given situation; the loosely defined concept of foveal vision is replaced by that of an effective glimpse aperture.

- $P_2$ is the probability that if a target is viewed foveally for one glimpse period it will, in the absence of noise (i.e., $S/N \gg 1$), be detected. $P_2$ is determined by psychophysical limits operating on the observed target size and contrast.

- $P_3$ is the probability that if a target is detected it will be recognized (again during a single glimpse and in the absence of noise). Recognition is usually (but not necessarily) accomplished on the basis of intrinsic shape without reliance on context.

We then write for $P_R$, the probability of target recognition,

$$P_R = P_1 \times P_2 \times P_3 . \qquad (1)$$

Inasmuch as the three steps described above are independent events, $P_2$ and $P_3$ (as defined) each representing a conditional probability under the one preceding it, the product formulation of Eq. (1) is obvious and rigorously correct. This is so despite the fact that the individual terms are not completely independent in the sense that they may be functions of some of the same variables (contrast, for example). This and certain other subtle interactions are discussed briefly in a later section. In the following paragraphs the nature of each of the three terms is examined in some detail, and a specific analytical expression is developed for each one.

### The Search Term

The first term, $P_1$, describes the search limitations. When searching from the air for a terrestrial object whose location is known only approximately, it becomes both possible and necessary to utilize foveal vision and to search fairly systematically. The maximum acuity of

foveal vision is a necessity, since there is always a need to find targets at the earliest possible moment during approach, and at long range either the apparent size or the available contrast or both may be marginal; few military targets really stand out. Foveal vision is also usually feasible, since only a limited area needs to be covered. The required area may be as much as the whole of an electro-optical display, but more commonly it is an area set by navigation errors and target location uncertainties, centered on a predicted or expected target location. Even under these conditions, however, search rates are extremely variable and almost intractable for the fundamental reason that pieces of terrain (not to mention possible targets) differ widely and almost defy quantification. Nevertheless, some bounds can be set.

It is well known that the eye moves in discrete steps, ordinarily with about three stops, called fixations, per second. (Actually an observer occasionally takes longer to examine certain points, but this does not affect very much the average search rates described below.) Our approach, therefore is to postulate that an experienced observer searches by moving an apparent aperture (essentially his foveal vision) in some fairly regular pattern over the area of interest, and furthermore that he adjusts his average interfixation distance, and hence the effective size of his scanning aperture and his overall search rate, in accordance with his *a priori* information on the size and contrast of the target or its image. Intuitively, one recognizes that an observer will scan the floor around him differently if he is looking for a pencil or an ant. Stated more formally, the observer estimates how far off his visual axis he will still have an adequate probability of detecting the expected image, and he automatically adjusts his search rate accordingly. A key concept, therefore, is the size of the effective scanning aperture--here called a glimpse aperture, $A_g$. This is a quantity that commonly ranges from 10 to 100 times the area of the target, $a_T$, but can sometimes vary between 1 and 1000 times $a_T$.

The reason for this huge spread is not just the observer's inability to predict the nature of the image or his own detection probabilities. It lies also in a second important factor--the structure, complexity, or "congestion" of the surrounding scene. The search for an ant mentioned above will also be quite different depending on whether the floor is covered with a nearly featureless linoleum or a textured and patterned rug. However, this "congestion" cannot be described solely by the two-dimensional spatial-frequency content in a scene. What really matters is the density of contrast points--the natural fixation centers for the eye--or other "confusion objects" that are present in the scene. The writer once experienced a striking example of many such false targets (natural decoys, as it were) while flying over the notorious Coso Range in California. This region contains scattered trees and bushes which appear very dark against the background of sandy soil or dried grass, as do the vehicles and "bridges" which were placed in the area as "targets". Almost every tree had to be examined to see whether or not it had straight sides before the true targets could be found. Indeed, tests there have produced some of the lowest target acquisition probabilities ever measured.

The kind of adaptive search rate described here, in which the observer automatically reacts to both the character of the scene and the (anticipated) nature of the target imbedded in that scene, has been advocated informally by this writer for several years. Only one independent reference to such a concept was found in the literature (Williams, 1966). Williams talks about target "conspicuity", which is measured by the rate at which a particular target can be successfully searched for in a particular field, and he points out that the commonly observed lack of dependence of target acquisition on display scale factor (within limits, and assuming no change in information content on the display) is another manifestation of observer adaptation.

A heuristic derivation of an expression for $P_1$ follows. If an area $A_s$ is to be searched, the number of glimpses (each of area $A_g = ka_T$) required to cover the area is $A_s/A_g$. The number of glimpses that are available in t sec, at 1/3 sec per glimpse, is 3t. With perfectly

systematic search, the probability of "looking at" the target (i.e., including it within a glimpse aperture) would be just the ratio of the available glimpses to the total number required, or $3t/(A_s/A_g)$. This would give $P_1$ the form of a linear ramp function with time. Real search is probably something between perfectly systematic and purely random, so that $P_1$ should have a form that lies between the ramp and an exponential rise. We conservatively adopt the latter and postulate

$$P_1 = 1 - e^{-K \times 3t/(A_s/ka_T)},$$

where K is a constant and k is a parameter related to scene congestion.

The supporting experimental evidence and the evaluation of the constants can be found in the original reference (Bailey, 1970). However, one more idea must be introduced. Since the value of k, the number of target areas in a glimpse aperture, is usually found to lie in the range 10 to 100, we substitute 100/G for k, where G is a measure of the congestion in a scene. Formally, it is the number of fixation centers within the nominal glimpse aperture of 100 $a_T$; in practice, it is nothing more than an estimate by the observer of the congestion of a particular scene relative to his experience, and he is asked to assign a number which usually falls between 1 and 10. Although this estimate may be quite crude, such as no better than within a factor of two, it is believed to be far better than ignoring the problem altogether. Further studies are needed to refine this troublesome point.

Accordingly, we propose the following expression for $P_1$:

$$P_1 = 1 - e^{-[(700/G)(a_T/A_s)t]}. \qquad (2)$$

The Contrast Term

The second term, $P_2$, has to do with the basic process of contrast detection by the human visual system. Blackwell's classical experiments provide the fundamental data here, yielding curves of threshold contrast (50 percent detection probability) versus size of circular discs under various levels of ambient illumination. These are commonly called "demand" contrast functions. However, there is a good deal of evidence that the best (i.e., lowest) threshold values obtained by Blackwell must be adjusted upward substantially for application to the practical situations discussed in this paper. Ignoring some of the details, which can be found in the original reference (Bailey, 1970), it is proposed that the shape of the "demand" curve of threshold contrast $C_T$ versus angular subtense $\alpha$ in minutes of arc (min) be taken from the average of the two best-known sources of 1/3-sec data, and that this curve be adjusted upward by a factor of about 5.5 in contrast--or that 0.75 be added to log contrast. Since, in the absence of bright lights or specular glint, target contrasts (using the absolute value definition) greater than unity are rarely observed through the real atmosphere, and even less frequently on military targets, the resulting curve on log-log paper can be approximated by the hyperbola

$$(\log C_T + a)(\log \alpha + 0.5) = 1. \qquad (3)$$

This simplification is often convenient and usually adequate, but, whenever contrasts greater than unity are important (for example, on certain electro-optical displays), a more accurate curve with an asymptotic slope of -1/2 should be used.

The probability of detection, $P_2$, at the threshold contrast is, by definition, 50 percent. The probability of detection for other values of observed contrast, C, has been shown to depend only on the ratio $C/C_T$ and to have the form of the cumulative normal distribution with $P_2 = 0.9$ for $C/C_T = 1.5$. This is equivalent to setting the value of the Gaussian standard deviation equal to 0.39, and it indicates that on the average Blackwell's subjects chose to operate at a false-alarm rate of about 1/200, corresponding to an S/N of roughly 2.6:1. Accordingly, we write

$$P_2 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\{[(C/C_T)-1]/0.39\}} e^{-u^2/2}\, du. \qquad (4)$$

## The Resolution Term

The third term, $P_3$, has to do with the more subjective act of deciding what particular image forms represent in the real world. But, since we are primarily concerned with shape recognition of known or briefed objects as distinct from the interpretation of unfamiliar imagery, the problem can be reduced to the visibility--or detectability in the sense of the previous section--of sufficient geometrical detail for shapes to be compared with memory and thereby recognized. The concept of "sufficient" detail might lead one into the morass of "critical details"--those unique features that permit various classes of objects to be distinguished from one another. However, when all portions of an image are equally detectable so that the whole shape is either visible or not, Johnson of NVL has demonstrated the remarkable fact that, for a variety of military objects*, a single parameter--namely $N_r$, the number of resolution cells contained in the shortest dimension across a target--is all that is required to describe what constitutes "sufficient" detail for detection or for recognition. He found values of $N_r$ between 3.3 and 4.8, or 4.0 $\pm$ 20 percent, for high-confidence recognition. Other experimenters have confirmed this simplification and derived values for $N_r$ close to Johnson's or slightly larger. We adopt a conservative value and write

$$P_3 = 1 - e^{-[(N_r/2)-1]^2} \qquad N_r \geq 2 \qquad (5)$$

$$= 0 \qquad N_r < 2$$

which makes $P_3 \approx 0.9$ when $N_r = 5$.

It is important to emphasize the meaning of $N_r$. As previously defined, it is the number of resolution cells contained in the minimum dimension (e.g., width or height) of the projected image of an object to be recognized. In the present context, "resolution cells" means independently detectable spots--the subject of the previous discussion. The proper procedure is to first calculate, from Eq. (3), the size of the smallest spot that can be seen--at the contrast level with which the target is presented to the observer. Then, the number of these detectable spots contained in the shortest dimension of the target image gives the value of $N_r$.

It was implied above that recognition in unfamiliar situations may be much more complicated, and far more difficult to predict, than the mere detection of shape details. An extreme example might be the classical one of the photo-interpreters searching for completely unknown elements of the Peenemunde launching areas during World War II. No attempt is made to extend this model to cover such cases. It should also be mentioned, however, that under certain other circumstances recognition may be very much easier than this model would predict. Consider the approach of unauthorized aircraft, or the presence of vehicles along a road in enemy territory. Both are cases in which the mere detection of objects might be sufficient to justify the decision, "There is a target!". These cases can be handled by assigning artifically high values to $P_3$ (when the prior information so justifies), thus effectively equating detection as given by $P_2$ to recognition. This point is discussed further below. Our model of $P_3$ covers the more common intermediate cases in which shape provides the primary criterion for recognition.

## Discussion

Equations (2) to (5), combined as indicated by Eq. (1), constitute the proposed model of a human observer. The fundamental concepts and the basic product formulation are explained at the beginning of this paper. The result, after analyzing each of the three terms, is an expression for the probability of recognition as a function of several observable quantities--the apparent size and contrast of the target, and the required search rate and the false-target density in the scene.

---

*One should probably add "in a military context". The amount of detail required to distinguish a truck from an ox-cart is far less than that required to discriminate between various truck models; but the simple separation of objects into classes is usually sufficient for designating targets.

An important and useful property of the model is the separation of variables that has been achieved. Each of the terms is expressed as a function of a rather small number of input parameters, and target size and contrast are the only parameters which enter into more than one term. This rather significant simplification arises from a careful consideration of the consequences of the product formulation and a detailed evaluation of each of the terms over only the ranges of the input variables for which that term is controlling or otherwise of interest.

For example, the model is not applicable to a target that is so isolated or whose contrast is so high (relative to the background clutter) that it can easily be seen with peripheral vision*, since in that case the search rate can be very much faster than postulated in Eq. (2) and $P_1$ will still be very high. But then $P_2$ will also be very high (essentially unity), and the problem is almost trivial. The search model assumes only that target contrast is *not* that high, so that fairly systematic and fine-grained search must be carried out. In fact, the actual search rate employed by an observer is determined by some sort of average false-target density over the scene. If the actual contrast of a specific target against its contiguous background turns out to be less than sufficient for recognition to take place during a single properly-directed glimpse, this fact will show up in $P_2$ and $P_3$, which will correctly reduce the value of $P_R$.

The relationship between the conditional probabilities, $P_2$ and $P_3$, can be discussed in a similar way. While it is clear that they are intimately related, they are separated, with further separation of variables, for several reasons. Ordinarily, detection not only precedes but also dominates shape recognition. That is, unless $P_2$ is rather high, there is probably no point in even calculating $P_3$, since $P_R$ will be too low to justify the sortie. When $P_2$ is high, then $P_3$ controls. On the other hand, as has been mentioned, there are cases in which *a priori* or contextual information may suffice to obviate the need for shape recognition per se. In such cases--boats on a river or trucks on a road, for example--$P_3$ can be ignored (i.e., set to unity without regard for Eq. (5)) and $P_2$ will control. By keeping the two terms separate, model flexibility is preserved. Further arguments for this separation revolve around the role of resolution. First, as a practical matter, most man-made sensor systems are resolution-limited, since resolution always costs something. (This is true at least of systems whose displays are properly designed.) Accordingly, the sometimes-difficult calculation of system MTF need be applied only once (namely, when it is most critical) in the shape-recognition term. More importantly, there are many cases with multiscaled or zoom-capable systems in which the combination of *a priori* information and required search area may make a two-step identification of the target desirable. In such cases an initial and tentative detection on a wide field of view is followed and confirmed (or denied) by shape recognition on a magnified image. At the first step $P_2$ controls, but $P_R$ is incomplete; at the second step $P_3$ controls.

The product of the three terms provides a viable model for a wide variety of circumstances; it can be used in predicting the capabilities of a broad class of manned systems, since it deals only with the observer and the information presented to him, whether this be directly to his unaided eyes or through optical aids, or with sophisticated artificial sensors provided that the signal-to-noise ratio is high.

## APPLICATION TO FLIGHT TOWARD A TARGET

The foregoing basic model really applies only to static situations, as when looking from a hovering helicopter. In normal flight approaching a target, the time constraints enter in a somewhat different manner.

Firstly, it is observed that the product $P_2P_3$ gives a single-glimpse probability of recognition as a function of the distance from the target. The distance, together with the flight path and altitude of course, determine the proper projection,

*This is essentially what is achieved by multispectral cueing or by MTI radar, as indicated on page 8 of the original reference (Bailey, 1970).

the angular subtense, and the apparent contrast (through the atmosphere) of the target. Secondly, repeated glimpses in the same direction (i.e., fixed on the same contrast point) are not independent in the usual probablistic sense and do not, per se, increase $P_R$ significantly. What does change with repeated looks in normal flight is that the observer is getting closer to the target. Thus the $P_R$ vs D curve with $P_1 = 1$, which may look like a cumulative probability curve, really is not that; it is simply a plot of the increasing probability of recognition as one approaches a target--on the assumption that no search at all is required.

The need to search for a target can only delay its acquisition, or, saying it differently, a given probability of recognition will be realized at some shorter range. To determine the amount the original curve must be shifted toward later (i.e., smaller) values of range, one must utilize the concept of searching over a specific area, $A_s$--such as a SAM-site pattern, or the area within a revetment, or the area enclosed by a reticle that is adjusted to indicate the navigational error, to mention a few examples. If the search over this area is thoroughly systematic, which it might be since the search is completely structured, then the maximum delay, $\tau$, would be the time taken to search the indicated area (given by $A_s/3A_g = A_s/3ka_T = A_sG/300a_T$). On the average the delay would be close to half that quantity, and the corresponding range shift would then be $V\tau/2$, where V is the approach velocity. (When the search area is not well defined, as in locating the first and grossest checkpoint, an arbitrary expression of the form $\tau$ (sec) = 1 + V(kn)/200 has been used.)

If the search were indeed ideally systematic, the time delays would lie between zero and twice the average value; and indeed this kind of spread in field test data must be anticipated. With a search pattern that is less than ideal, the $P_R$ vs D curve will be both displaced and flattened, with the result that the probabilities calculated by the method described will be a little optimistic at the high side (i.e., close to the target).

## APPLICATION TO MOVING TARGETS

In 1970, Rand conducted some simple simulation experiments (Dugas, 1971; Petersen, 1971) on the detection of stationary and moving targets. The data consisted of the times taken by several observers to search a TV screen and find a simple, electronically generated, rectangular target against a background that was variously a piece of felt, an artificial grid pattern, or scenes from aerial photography. The exponential time-dependence was confirmed, and the effect of target motion could be well accounted for by introducing into the exponent of the $P_1$ term a factor $(1 + 0.45\ V^2)$, where V is the target velocity in degrees/second subtended at the observer's eye. There was some indication that part of the effect in moving objects being "easier" to detect is due to the changing contrast as the object passes over a complex background--analogous to the detection of flashing lights. However, this effect could not be quantitatively separated out in these experiments.

## CONCLUDING REMARKS CONCERNING NOISE

The foregoing description has ignored noise altogether. Historically, I used to think that this model could be equally applicable to an observer viewing an E-O display, except that his performance would necessarily be degraded by the insertion of electrical noise. I therefore multiplied the $P_1P_2P_3$ product by a fourth term which was given by

$$\eta = 1 - e^{-[(S/N)-1]} \qquad S/N \geq 1$$
$$= 0 \qquad S/N < 1 .$$

This form has been used at Rand, and by other people too (I'm told), to make performance predictions that seem to be reasonable. It is certainly better than nothing; but it is now clear that the concept of "perceived signal-to-noise" described here this morning, and the procedures that were given for applying it, constitute a much better approach to understanding recognition through noise. The procedures may appear to some to be a little arbitrary or *ad hoc*-ish, but they do work. Perhaps a more fundamental (or at least a more satisfying) theoretical approach can be evolved building on Ory's

work, mentioned earlier in this paper. Perhaps someone in this audience will do just that in the very near future.

## REFERENCES

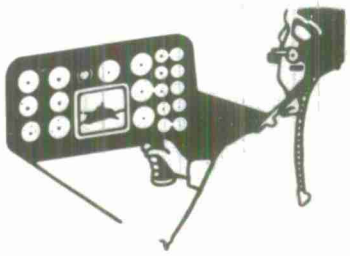

Bailey, H. H., *Target Detection Through Visual Recognition - A Quantitative Model*, RM-6158-PR, The Rand Corporation, February 1970; also appeared with classified applications in Journal of Defense Research, Vol. 3B, p. 54, 1971.

Dugas, D. J. and H. E. Peterson, *An Experimental Investigation of the Effect of Target Motion on Visual Detection*, R-614-PR, The Rand Corporation, February 1971.

Ory, H. A., *Statistical Detection Theory of Threshold Visual Performance*, RM-5992-PR, The Rand Corporation, September 1969.

Peterson, H. E. and D. J. Dugas, *The Relative Importance of Contrast and Motion in Visual Target Detection*, RM-688-PR, The Rand Corporation, March 1971.

Williams, L. G., *Target Conspicuity and Visual Search*, Human Factors, Vol. 8, p. 80, February 1966.



# THE STATE-OF-THE-ART IN FLIR TARGET ACQUISITION MODELLING

by

Antonio J. Mendez and Melvin Freitag
Martin Marietta Aerospace
Orlando, Florida

OFFICE OF NAVAL RESEARCH
TARGET ACQUISITION SYMPOSIUM
NAVAL TRAINING CENTER, ORLANDO, FLORIDA / 14,15,16 NOVEMBER 1972

# THE STATE-OF-THE-ART IN FLIR TARGET ACQUISITION MODELLING

Antonio J. Mendez and Melvin Freitag
Martin Marietta Aerospace
Orlando, Florida

## ABSTRACT

A survey and critique is made of nine analytical models which have been developed or used recently to describe the performance of thermal imaging sensors as target acquisition devices. The models are those developed by Bailey (Rand), Bailey/Martin (Martin Marietta), the MARSAM (Honeywell), the one developed by NAFI, the classical model of Texas Instruments,the probabilistic model of Texas Instruments, the FLAG (Hughes) model, the one developed by NVL (Ft. Belvior), and the one developed by EOS. Mention is also made of current modelling efforts at NADC to model ship classification with FLIR sensors.

A brief discussion is then given of the critical physical operations and functions associated with target acquisition, as well as the more primitive phenomena from which these operations and functions are structured.

The critical operations are defined in terms of (presumably) measurable parameters. The state-of-the-art of the modelling depends on the understanding and empirical status of these parameters. An examination of this status is given.

## INTRODUCTION

Attention has focussed in recent years on thermal imaging sensors because of their impressive operational target acquisition performance. These sensors, usually of the forward-looking infrared (FLIR) kind, reached an engineering maturity during the middle and late sixties. The engineering and operational evolution have progressed at a rate that easily outdistanced the analytical and phenomenological understanding of these devices with respect to target acquisition (an excellent historical and functional treatment of FLIRs is given in Lloyd's Thermal Imaging Systems, to be published[1]). This situation has led to efforts to rigorously understand the sensor parameters which influenced target acquisition and to standardize their definition and specification (Biberman's FLIR Specification Guide [2].)

There is at this time by no means a universally accepted approach to FLIR target acquisition modelling. This will be brought out by this paper, which will survey several models which have appeared recently in the literature or which have been used recently in documents describing relationships between sensor parameters and target acquisition performance.

This paper will also describe the scenic, physical, sensor, and pychophysical relationships which seem to impact target acquisition performance. The paper will close with comments on the state of understanding these relationships, and descriptions of some recent work towards improving the understanding.

## TYPICAL MODELLING APPROACHES

Most target acquisition models have the structure of a cascade or nesting of conditional probabilities. The various conditional probabilities have, in general, different content and meaning in the various models. Some of the conditional probabilities in a given model are functions of the same parameter, so that it has been suggested that the usual product form should be replaced in part by convolutions.[3]

The individual conditional probabilities invariably are derived from an analysis of independent experiments.[4] To date target acquisition measurements have not been carried out which allow a proper evaluation of the relative importance of the phenomena described by the individual relations. Because of this, the transition among the various phenomena in the course of target acquisition is not well modelled, hence the reliance on simple statements such as nesting of conditional probabilities.

Tables I through III describe functionally the MARSAM II, Bailey and Bailey/Martin models of target acquisition. The tables illustrate the nesting approach, as well as the different content in the elements of the nesting.

Most models derive the probabilities of target detection, recognition and identification in separate computations. Assumptions are then made about the transition from detection to recognition (for example) in order to compute system performance (with auxiliary mission data), such as probability of first pass attack. Appropriate measurements have not been carried out and/or analyzed to allow the correct mission characterization of the transition from detection to recognition and identification.

Some of the models are designed to describe the sensor performance in the sense of image quality, having in mind perhaps the requirements of a photo-interpreter (e.g., MARSAM II). These models tend to be quite analytical. Other models are designed to associate the sensor's primary specifications with real and stylized target acquisition (e.g.,Bailey/Martin). These models tend to be phenomenological. Hybrid (analytical/phenomenological) combination also exist (in fact, most models).

## TARGET ACQUISITION MODEL SURVEY

Tables IV through XII depict the essential contents of nine current FLIR target acquisition models. Two other modelling efforts are not depicted but are of importance: NADC's (Moser)[6] analysis of ship classification via FLIR imagery and APG's (Pibble)[7] modelling of thermal signatures and their sampling (for IR seeker optimization).

The tables indicate that most models ignore the search portion of target acquisition. One of the models (Bailey/Martin) has included search as an efficiency factor in which the human scan tendencies, the scan geometry (imposed by the mission and geography, and prebriefing), the time available to scan ( imposed by geometry, V/h and FOV) and the scene congestion (number of potential false targets) determine the dwell time on the true targets. All models which are philosophically similar to the Bailey model are compatible with such a search term. The identification of clutter with false alarms is not made exact in any of the models.

The tables also show that several of the more developed models are derivatives of the Bailey approach, indicating a certain aesthetic and intuitive appeal of that approach.

Theoreticians in this discipline exhibit extreme differences in the description of the detectability and resolution criteria upon which their modelling is based. Some of the models make target detectability a function of satisfying the Blackwell criteria; others (notably NVL) make it strictly a function of signal to noise.

With respect to resolution criteria, some of the models describe pattern recognition as a certain number of "resolvable spots" for a given level of task difficulty (detection, recognition, identification). Among these are the Bailey, NVL, NADC models. Other models, notably the Bailey/Martin, take a somewhat pragmatic approach and associate with each class of tasks and targets an empirical "bar target equivalent." Models which mix the resolvable spots and bar target equivalents notions are not well thought out.

The resolvable spots approach is analytically akin to the Schade measures of image quality such as $N_e$ [15]. Thus, if and when an explicit FLIR model is based on MTFA, [16] it will fall in this category.

TABLE I

MARSAM[5] Mathematical Model of Target Acquisition

Basic Definitions

| | |
|---|---|
| $P_{los}$ | The probability of having a line of sight to the target |
| $P^*_{d1}$ | Conditional probability that the element size and contrast are sufficient for detection. |
| $P^*_{d2}$ | Conditional probability that a displayed element of sufficient size and contrast will be fixated by an observer within a specified time. |
| $P^*_{d3}$ | Conditional probability that an observer-viewed element of sufficient size and contrast will be correctly discriminated in the presence of confusing objects. |
| $P^*_{d4}$ | Conditional probability of sufficient signal to noise for detection. |
| $P_d$ and $P_r$ | Computed on the assumed statistical independence of $P_{d1}$, $P_{d4}$ and $P_r$ with respect to each other and with respect to $P_{d2}$ and $P_{d3}$. Effects of confusing objects are not independent of time and the search mode. $P_{d2}$ is expressed as a function of $P_{d3}$. |

Basic equations:

$$P_d = P_{los} \circ P^*_{d1} \circ P^*_{d2}(P^*_{d3}) \circ P^*_{d4}$$

$$P_r = P_d \circ P_r^* \quad = \quad \text{conditional probability of target-element recognition}$$

(Based on Defense Research Corporation model from Stathacopoulos)[29]

TABLE II

Bailey Model of Target Detection Through Visual Recognition[4]

Basic Definitions

1. $P_1$ is the probability that an observer, searching an area known to contain a target, looks for a specified glimpse time in the direction of the target with his foveal vision.

2. $P_2$ is the probability that if a target is viewed foveally for one glimpse period it will, in the absence of noise, be detected.

3. $P_3$ is the probability that if a target is detected it will be recognized (again during a single glimpse and in the absence of noise).

4. Quantity n is an overall degradation factor arising from any noise in the image that is viewed by the observer.

BASIC EQUATION

$$P_R = P_1 \circ P_2 \circ P_3 \cdot n$$

TABLE III

Bailey/Martin Marietta Mathematical . el of Target Acquisition[3]

Basic Definitic ıs

$P_1$ = The conditional probability of fixating the target with the operator's glimpse aperture (fovea) within the available search time which is determined by the sensor's FOV, the mission parameters and the scene clutter, i.e., false target density; the FOV and mission parameters determine the rate of change of area to be searched; prebriefing and scene cues are introduced in the geometric definition of the search (spot, gross or detailed linear, gross or detailed area, and pseudo random);

$P_2$ = The conditional probability of having the required contrast/target size relationship to mediate target detection given that the target is in the operator's glimpse aperture; if the scene statistics (contrast distribution in the scene) is known, the $P_2$ as defined is convolved with the scene statistics to give the effective $P_2$;

$P_3$ = The conditional probability of target pattern recognition reduced to a bar target equivalent; targets and tasks (detection, recognition and identification) are ascribed different bar target equivalents, and these are defined empirically; bar target equivalents are not defined for some targets and tasks (e.g., ship classification);

$P_4$ = The conditional probability that there is sufficient (determined empirically) displayed signal to noise ratio (SNRD) to mediate the bar target equivalent.

The basic equations are:

$$P_D = P_1 \cdot P_2 \cdot P_3(D) \cdot P_4(D),$$

$$P_R = P_3(R) \cdot P_4(R),$$

$$P_I = P_3(I)\ P_4(I),$$

for detection, recognition and identification, respectively. The symbols D,R,I represent the functional description of the bar target equivalents for detection, recognition, identification. Typically, these can vary with the type of sensor used.

TABLE IV

A Summary of the Honeywell Multiple Airborne Reconnaissance Sensory Assessment Model (MARSAM II)[5]

| | | |
|---|---|---|
| 1. | TYPE | Stored Program, Data Bank |
| 2. | PRIMARY PURPOSE | Assess Reconnaissance Sensor Systems |
| 3. | SENSOR TYPES | Frame Camera, Panoramic Camera, TV, Visual Observer, Vertical IR, FLIR, SLAR, FLR, ELINT |
| 4. | TERRAIN AND FOLIAGE MASKING | As submodel, computes probability of having a line of sight to target |
| 5. | HUMAN FACTORS SUBMODEL INPUT VARIABLES | Target element/background display Contrast Angular Element Size, Time Duration of Scene, # of Confusing Objects, SNRD |
| 6. | BASIC STRUCTURE | $P_d = P_{los} \circ P_{d1} \circ P_{d2}\{P_{d3}\} \circ P_{d4}$ <br> $P_r = P_d \circ P_r^*$ |
| 7. | SEARCH MODEL | Random search only, 5° glance at the eye (each glance independent), Based on fixation area, computes expected number of fixations. |
| 8. | RECOGNITION SUBMODEL | Based on Brainard's scan line requirements data, TV and IR = $N_t$ 6 scan lines for $P_r$=.5. |
| 9. | ATMOSPHERIC MODEL | Extensive weather model (5 types in store), uses atmospheric layers to compute transmittance, apparent path luminance computed as a function of backscatter and ground illumination. |
| 10. | TARGET-ELEMENT DISTINCTIVENESS | Computes target-element distinctiveness coefficient which accounts for cuing. |
| 11. | SENSOR MODEL BASIS | Limiting resolution of sensor projected on ground. |
| 12. | FLIR TRACKING MODE | Evaluates effectiveness of tracking or not tracking targets to enhance recognizability. |
| 13. | CONTRAST TERM | $P^*_{d1}$ based on Blackwell corrected by .75 log units (as recommended by Blackwell). Based on Contrast Ratio (Apparent vs contrast threshold). Corrected for brightness of background less than 10 ft. Lamberts. |

TABLE IV (Cont)

| | | |
|---|---|---|
| 14. | $P_r^*$ | Uses no. of visual acuity elements (for direct vision) across effective min. target-element dimension. More than 3.2 required. (arc min.) |
| 15. | PLATFORM MOTION | Computes platform motion limiting by computing forward motion compensation in TV model, image smear distance. Resultant combined effects evaluated by increasing ground resolution distance D. |
| 16. | SNR | Based on Schade's data, empirical fit 1.4 for P=.50, P=.90 at about 3.5. Uses Stathacopoulos for number of TV scan lines vs $P_r^*$. |
| 17. | DISPLAY MOVEMENT | Minimum threshold of $P_r$ for display target size as a function of angular rate. (TV and FLIR) uses Dugas for threshold function. Exponential increases in visual acuity threshold above 25 degr/sec. Same in visual observer model. |
| 18. | VISUAL OBSERVER MODEL | Computes circular search area offset from flight direction, uses minimum search range, maximum search range to obtain $P_{los}$ and $A_{eg}$ for far middle and near points for averaging. |
| 19. | BACKGROUND AND TARGET REFLECTIVITY (or equivalent IR characteristic) | Background reflectivity input used in confusing objects submodel and background brightness computation. Also, target-element reflectivity in observer and camera models. |
| 20. | CAMOUFLAGE | Camouflage contrast degradation factor reduces contrast in observer TV model and camera models. |
| 21. | TARGET TYPES | Area and linear types (radial extent of area target) |
| 22. | DISPLAY TYPE | Stationary moving or track modes possible in human factors display submodel. |
| 23. | MULTI-SENSOR IMAGERY | Hard-copy only, uses weighting coefficients for 4 spectral regions for each of D/R/I, 4 spectral regions applied to 15 types of targets. Uses Inter-Service Multi-Sensor Imagery and Data Storage and Retrieval Systems data. |
| 24. | THERMAL BASIS | Power received from target and background and calculated NEP. |

TABLE V

A Summary of the Bailey/Martin Model of Target Detection Through Visual Recognition[3]

| | | |
|---|---|---|
| 1. | TYPE | Calculational (atmospheric data stored). |
| 2. | PRIMARY PURPOSE | To describe quantitatively the capabilities and limitations of a human observer with respect to target acquisition using an imaging sensor. |
| 3. | SENSOR TYPES | CRT displayed TV, RGTV or FLIR imagery. |
| 4. | TERRAIN AND FOLIAGE MASKING | Explicitly included in range gated TV (RGTV) routine; FOV projection (footprint) included in TV, RGTV, FLIR computer routines. |
| 5. | HUMAN FACTORS SUBMODEL INPUT VARIABLES | Psychophysical model. Search, detection of contrast, recognition of shapes; all conditional probabilities and a degradation factor, based on SNRD. |
| 6. | BASIC STRUCTURE | $P_{d,r,i} = P_1 \circ P_2 \circ P_3 \circ P_4$, with $P_1 = P_2 = 1.0$ for $P_r$ & $P_i$ |
| 7. | SEARCH MODEL | Based on glimpse, moving aperture (adjustable in size) for search model. Ranges from random to organized search depending on mission parameters. |
| 8. | RECOGNITION SUBMODEL | Based on MAFLIR data (FLIR model only) and NVL data, hence phenomenological. |
| 9. | ATMOSPHERIC MODEL | Barhydt (Hughes) humidity and temperature algorithm used to compute equivalent water per 1000 ft.; transmittance curves entered to determine degradation coefficient; altitude and slant range entered; atmospheric layering and scattering currently not entered. |
| 10. | TARGET-ELEMENT DISTINCTIVENESS | Uses search type to change number of glimpses required; uses clutter to determine search time available per FOV. |

TABLE V (Cont)

| | | |
|---|---|---|
| 11. | SENSOR MODEL BASIS | Thermal Signature and MRT of sensor, i.e., based on Johnson type bar target equivalents. |
| 12. | FLIR TRACKING MODE | No capability to simulate tracking mode. |
| 13. | CONTRAST TERM | Based on Blackwell corrected according to Davies. Also adds .75 to log contrast. Computes $C/C_t$ $(P_{90})$=1.5. |
| 14. | $P^*_r$ | U. S. Army NVL and MAFLIR data for $P_3$=.9; invokes Johnson notion of bar target equivalent. |
| 15. | PLATFORM MOTION | None (Motional MTF included as degradation to sensor MTF and MRT). |
| 16. | SNR | Separate $P_4$ Term; uses MRT to compute SNRD for bar target equivalent (separate computation for detection, recognition, identification). |
| 17. | DISPLAY MOVEMENT | None (as originally formulated); can be included as a Motional MTF term. |
| 18. | VISUAL OBSERVER MODEL | Built into all conditional probabilities (empirical data). |
| 19. | BACKGROUND AND TARGET REFLECTIVITY (or equivalent IR characteristic) | Thermal signature inputted where available; otherwise strong dependence on MAFLIR measurements. |
| 20. | CAMOUFLAGE | Tacitly included in clutter and thermal contrast. |
| 21. | TARGET TYPES | None (uses critical dimension of target and bar target equivalent; other notions required for complex and maritime targets). |
| 22. | DISPLAY TYPE | None; CRT tacitly assumed. |
| 23. | MULTI-SENSOR IMAGERY | None; sensor models are independent. |
| 24. | THERMAL BASIS | MRT and target and background thermal signature (must be modified for maritime targets); atmospheric attenuation included. |

TABLE VI

NAFI Imaging Infrared Systems Math. Model[8]

1. No search probability term.

2. Calculates NETD, NEFE, and S/N (theoretical) (NEFE = noise equivalent fractional emissivity).

3. Computes angle subtended at the sensor by the target (calculated by geometry of situation and compares its value with the angle subtended by a single line of resolution).

4. Does not consider mission parameters, computes range-altitude performance envelopes.

5. Based on MTF including image motion.

6. Criteria: 1.5-2.5 lines for detection
   5-10 recognition
   10-16 designation (identification or classification)

7. Atmospheric model: (Townsend's model) 3-5 Microns only.

8. Non-probability organized program.

9. Major Basis: MTFs

10. Delta T of 10 degs. assumed (or target) S/N of 2 required for detection.

11. For Scanning Systems only.

TABLE VII

TI Probabilistic Model of the Target Acquisition Process for FLIR[9]

1. Atmospheric model: Wyatt, Stull and Plass, based on molecular absorption by water, and $CO_2$ gas and particle scattering by haze and/or light fog particles.

2. Based on S/N and MTF theory.

3. Based on Bailey's formulation.

4. Human operator model is based on the product of 3 terms; target contrast term based on Stathacopoulos et al and relates p of seeing to threshold contrast as a function of avg. target angular subtense at eye. Resolution term based on method of Stringham for detection and Bailey for recognition. Displayed S/N term based on Westinghouse and Hillman.

5. System model in that weapon delivery considerations are included.

6. Modulated with AN/AAD-4 FLIR flight tests (Ref. 22).

7. No search term in evidence; found search time in a cluttered background difficult to predict.

TABLE VIII

A Mathematical Model for Predicting the Probability of Detecting Target Against Various Background Utilizing a FLIR System. (TI Classical Model)[10]

1. Compares resolution element covering target with a resolution element of background radiation.

2. Classical model uses NEP, NET, S/N, D*.

3. Only considers target detection. No human factors element and display unit information included.

4. Single channel only.

5. Excludes atmospheric scattering and reflection. Ozone absorption in atmosphere ignored.

6. Atmospheric transmission based on Larmore/McGee and Larmore/Passman. Absorption effects of $CO_2$ and $H_2O$ considered.

7. Considers the oblique viewing effect.

8. Statistical analysis of the mean square noise fluctuation using Gaussian distribution. Transforms probabilities into a contrast scale to determine target/background relative contrast.

TABLE IX

Flag (Forward-Looking Air-to-Ground) Model of Hughes[11]

1. Includes human vision, IR and TV sensors.

2. A/C flight at constant velocity, altitude and offset toward the target; turns toward target if there is a navigational offset, includes time delays for mission geometry, and weapon preparation time.

3. Uses scanning angle limits, probabilities thresholds, sensor selector and target size.

4. Detection in terms of $C/C_1$ relative contrast.

5. Presented area of target computed in geometry subroutine.

6. Subtense, angle (alpha) computed from sensor, viewing parameters.

7. Computes line number included in the image subtense at the sensor input.

8. For recognition, uses resolution elements across target required for .50 recognition. Normalized to alpha, (target subtense on display). Uses eyeball resolution as a function of adaptation level.

9. Atmospheric and weather model included (considers relative humidity and ambient temperature and two weather states).

10. Includes emmissivity of target and background and uses delta T in IR model.

TABLE X

NVL (Ft. Belvoir) Thermal Imaging Systems Model[12]

1. TYPE: Calculational

2. PURPOSE: Input to meaningful trade-off analysis to help select devices to perform specific military tasks. To predict target detection-recognition ranges.

3. BASIS: Signal to noise/resolution model based on decision theoretical considerations. (Based on NET rather than SNRD)

4. BASIC EQUATION: A triple integration (two spatial and one temporal) of a Fourier transform of the signal after atmospheric attenuation, a Fourier transform of the signal after device and eyeball filtering and a "complex function of the signal and system parameters," divided by the noise power spectrum at the decision point. Detection probability is determined by a S/N threshold function.

5. HUMAN FACTORS: Enters into determination of the required S/N ratio.

6. ATMOSPHERIC MODEL: Hughes work based on Deirmendjian.

7. HARDWARE PERFORMANCE BASIS: System blur spread function (uses measured device MTF when possible).

8. TARGET DIMENSION OF SIGNIFICANCE: Area (rather than minimum dimension).

9. RECOGNITION CRITERIA: Function of the number of effective resolution areas (blur areas or resolvable spots) within the target.

10. NOISE COMPUTATION: From detector and display plus noise power spectral density of eye.

11. MODEL SHORTCOMINGS: Cannot handle the effects of filtered noise. Simple model that does not predict field performance (does not include enough operational and mission factors). Needs expansion to handle camouflage, non uniform effects, motion and vibration effects. No search model, although one is planned. No provision for clutter.

12. MODEL VALIDATION: Validated against Warren Grove field tests (predicts fairly well).

STRONGPOINTS: Good cost-effectiveness and cost trade-off analyses are possible utilizing NET as a basis. Very simple and straightforward mathematically. Can be expanded into comprehensive model by addition of subroutines for terrain, climate, weather, etc. Cheap and economical to implement in computer programming.

The NVL MODEL as detailed in the reference appears to be very promising as a simplified system model of target acquisition using thermal imaging systems. It needs further expansion to include more of the facts of thermal imaging systems in target acquisition tasks. As noted on the summary sheet, these shortcomings are in the area of better utilization of target signature data, scene clutter search and cuing factors. The variable set at present is believed to be too small to be useful where real-time operational systems and real eyeballs are involved. SNRD as an observable is not used as a definitive term in the probability determination. It is not clear what decision factors are entered into the decision function discussed in the report.[10] What is the decision criteria function; how is this determined; how does this enter into the final decision whether a target is there and what kind of target it is (recognition) all need clarification. It appears to take care of the recognition task adequately but not the detection task although the former is believed to follow sequentially in a decision scheme.

TABLE XI

Bailey Model of Target Detection Through Visual Recognition[13]

BASIC EQUATION

$$P_R = P_1 \circ P_2 \circ P_3 \cdot n$$

In applying his model to the case of locating vans contained in a mobile SAM site with a FLIR, Bailey makes several enlightening assumptions and simplifications in the quantification of his conditional probabilities (see Table II):

1. $P_1$ and $P_2$ are assumed to be unity since the effects of search are incorporated into $P_3$ and the display is assumed to be "ideal", that is, only signal to noise limited. He assumes no search is required because the target is easily located in a circumscribed area (SAM site) and because the display provides sufficient contrast, scaling, brightness of display, etc.

2. $$P_3 = 1 - \exp\left[-\left(\frac{L_{min}}{.4\,\alpha(R_s + V \cdot t_d)} - 1\right)^2\right]; \quad N_r \geq 0.4$$

3. $$n = 1 - \exp\left[-\left(\frac{\Delta T \cdot T_{at}}{NET} - 1\right)\right]; \quad T_{at} = \exp\left[-\sigma(R_s + Vt_d)\right]$$

$N_r$ = the number of resolution cells contained in the shortest dimension across a target

where $L_{min}$ is the minimum dimension of the target (shortest projected)

V is the velocity of the aircraft

NET is the Noise Equivalent Temperature

$t_d$ is the correction factor for dynamic viewing condition (from static detection)

$R_s$ is the range of the sensor from the target (slant range)

$\Delta T$ is the effective temperature difference of the target compared to the scene temperature

$\sigma$ is the absorption coefficient of the atmosphere, dependent mostly on water content

$\alpha$ is the resolution of the FLIR

Note that Bailey uses the criteria of number of resolvable elements which can be interpreted as resolvable spots or bars of a stylized (Johnson) target in the smallest dimension of the target for recognition.

Note also that Bailey uses a detector channel signal to noise (NET) ratio rather than a SNRD in the noise degradation factor.

MODEL STRONGPOINTS - AND SHORTCOMINGS - The Bailey model is straightforward, empirical in basis and easy to comprehend. Some of the computations (like resolvable elements, as described in Reference 4) are complex, however. It also takes into consideration the psychophysics of target acquisition. The applications given in the reference indicate that the 4 basic components of the model are not meant to be used all the time; their use depends upon the sensor and mission characteristics. This allows ample room for the analyst to use his judgment as to the limiting factors to be used for all types of systems and differing operational missions. Note also that in giving an example, Bailey sidesteps having to include search parameters ($P_1$ & $P_2$ are assumed to be unity). Note also that he assumes an ideal display which also sidesteps knotty display problems.

TABLE XII

EOS Model of FLIR Target Acquisition[14]

1. Basis: MTF and Noise Equivalent Temperature (synthesized into SNRD) Predictions are based on Bar-space Effective Temperature Difference, Display average brightness and System AC Gain.

2. Uses the concept of Resolvable Temperature.

3. Purpose of the Model: to predict the outcome of an IR Sensor Flight Test program using bar-type targets.

4. Target Acquisition Model: based on Bailey's model. Basic target acquisition equations:

$$P_D = P(L) \cdot P(D/L)$$

$$P_R = P(L) \cdot P(D/L) \circ P(R/D)$$

$P(L)$ = probability of looking at the target

$P(D/L)$ = probability of detection given looking

$P(R/D)$ = probability of recognition given detection

This model does not address search since in the test validating it the operator knew both where to look and what he is looking for. Note that this is a conditional probability model. It does not consider mission or flight parameters directly. This model starts with the target on the display. $P_R$ is equated to bar-breakout (the recognition criteria) utilizing display signal-to-noise ratio for determining $P_{D/L}$.

$$P_{D/L} = \text{Probability } (S + \text{noise}) > 1.8 \cdot \text{sigma}$$

$\Delta T$ is defined as the temperature difference which provides a S/N (Perceived) = 1.8. Uses Johnson's criteria for determining the number of resolvable elements required to discriminate the shape of the target. Uses the number of elements on the target minimum dimension (a la Johnson).

$$P(R) = 1 - e^{-(N/4 - .5)^2}$$

N = # of elements on the target minimum dimension

5. Shortcoming of the model: This model was especially derived to predict the results of a test program having very restrictive techniques and objectives and hence is not easily expanded or applied to other test programs, operational conditions or to other types of targets.

6. Strong points: Considers the eye-brain portion of the target acquisition system and includes display parameters not included in many models.

Theoreticians who use the bar target equivalent notion express their reliance on the empirical relations first formulated by Johnson [17] that real target acquisition and stylized target resolution are related (in a basically not well understood manner). The Bailey/Martin formulation falls in this category.

Moser's analytical work[6] in the classification of ships has shown that, for this class of tasks, the bar target equivalent notion is definitely inapplicable and the resolvable spot notion definitely valid. This kind of discrimination between notions has not been derived for terrestrial, compact targets. Some pessimism about the generality of the bar target equivalent notion has been expressed by Rosell.[18]

Sufficient confidence has been expressed in bar target equivalents in the past that one of the primary laboratory measures of FLIR performance, the MRT, is based on the imaging of a stylized bar target.[19] Also, electro-optical field test facilities, such as the Navy's at Patuxent River, have similar large scale stylized bar targets. Models using the resolvable spot notion should properly develop a corresponding sensitivity-resolution relation to replace the MRT.

The appeal of the bar target equivalent approach is that FLIR system specification and design, laboratory and field bar breakout tests, and real target acquisition performance can be readily associated with each other.

All models which reference the target acquisition performance to the information presented on a CRT display rely on the notion of the SNRD. As basic as this parameter is to the modelling, it is usually presented as a derived rather than a measured quantity. This situation is unsatisfactory for a discipline attempting to describe observable phenomena. Some efforts to systematize the measurement of SNRD are being developed by Synder.[16,21]

## DISCUSSION

Target acquisition involves the transfer of information from a scene to an observer. The transfer is effected by the sensor-display combination. The sensor's information transfer to the display is described analytically by sensitivity-resolution relations (and the like), and the observer's appreciation of the information is described by phenomenological threshold relations (such as the Blackwell criteria).

All complete models of target acquisition must eventually be reducible to the same set of observable or measurable quantities and threshold relations. Because of this, the discussion of any one complete model is sufficient to identify weaknesses in the understanding of the phenomena of target acquisition.

The Bailey/Martin model is one that has attempted to be complete, in that search through pattern recognition is included. Thresholds of the eye, dwell time, contrast-magnification, sampling-magnification and signal to noise type relations form the bases for the functional form of the model. The structure of the model is described in Table III and the physical observables and threshold relations in each conditional probability are shown in Figure 1. The conditional probabilities are designed around the observables and functions which seem to be of primary importance in FLIR target acquisition.

The figure shows, underlined, the FLIR related observables and threshold relations which need to be better understood, measured, or modelled. A natural result of improving these is better FLIR target acquisition modelling.

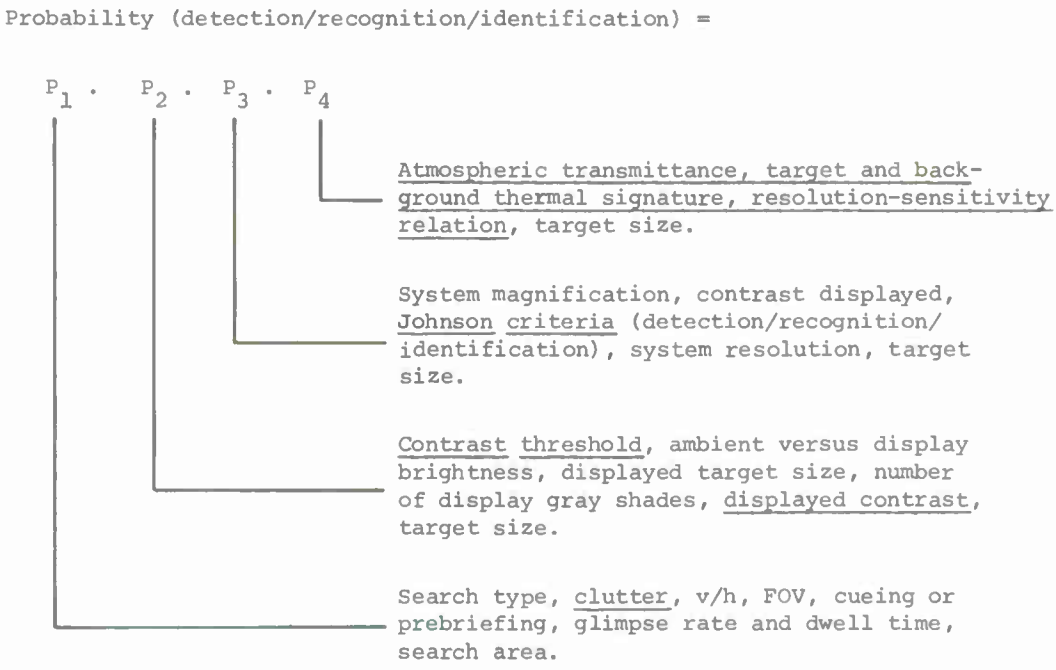



Figure 1. Description of a Typical FLIR target acquisition model in terms of the physical observables and threshold relations which govern the phenomena of target acquisition.

## RECENT CONTRIBUTIONS

This section describes recent contributions which impact the physical observables and threshold relations areas influencing FLIR target acquisition modelling as described in the previous section and Figure 1.

### Clutter

Clutter can be described as elements of the scene which have roughly the same contrast and dimensions as a real target in the scene. The importance of clutter is mainly in the search function. In recent work,[3] subjects were asked to identify frames from AN/AAQ-5 filmed imagery which could be quantitatively described as "cluttered," "moderately cluttered" and "uncluttered" (well defined target in an open field). In an effort to find a measure or discriminant for clutter, densitometer traces were made of these qualitatively characterized scenes. Fourier transforms and power spectral densities of the linescan were then made. It was found in this way that the power spectral density (in the spatial frequency domain) is a discriminant between cluttered and uncluttered scenes. This is illustrated in Figures 2 and 3. The discrimination is not yet quantized, but it is interesting to note that this discriminant is defined in the same domain that is used to describe the resolution and resolution-sensitivity relations of a sensor. It is hoped that, in the future, system studies of the resolution requirements of a mission will include, because of this discriminant, the impact of clutter imaged by the FLIR system.

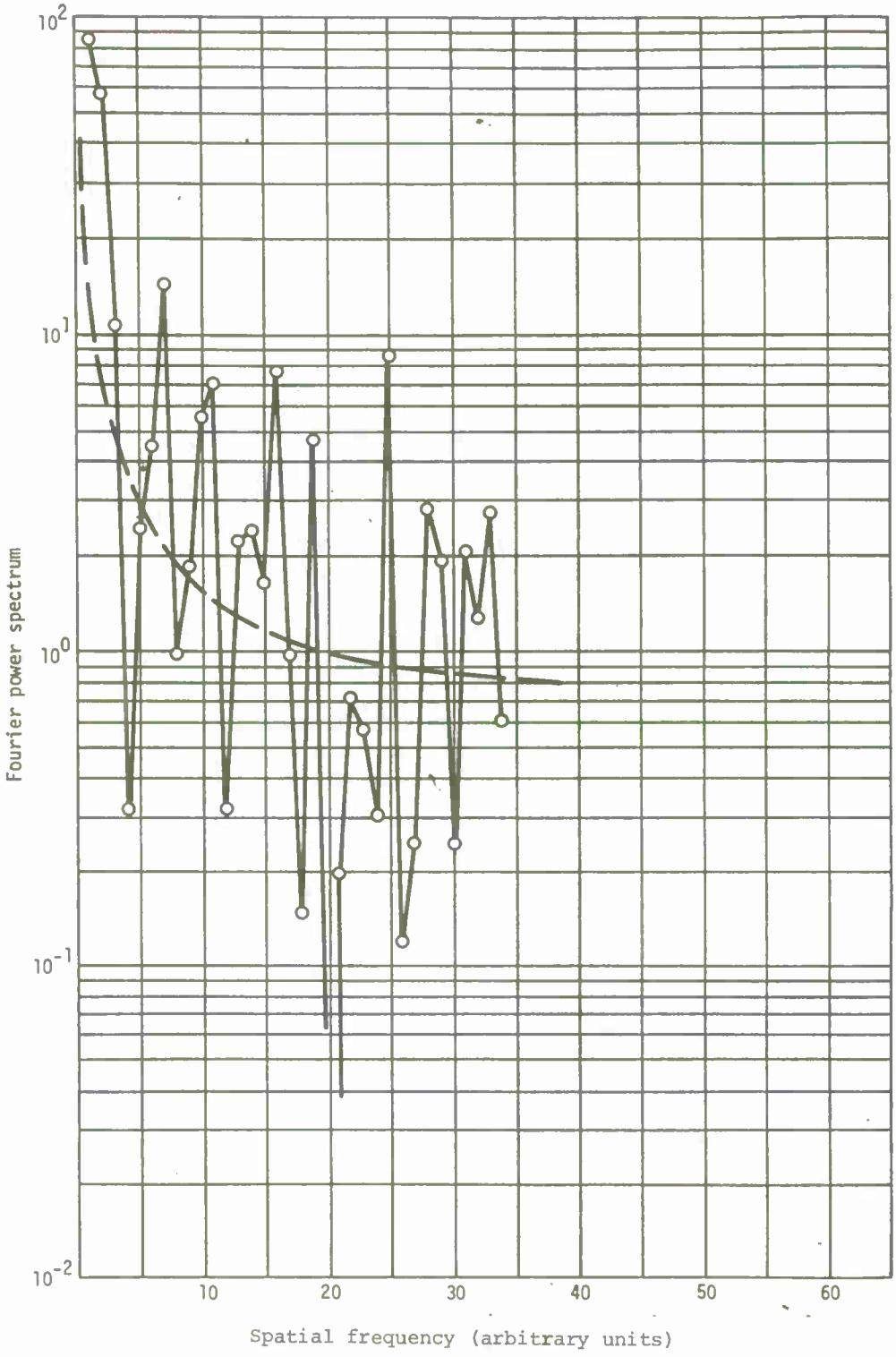



Figure 2. Power Spectrum Plot for a Typical Cluttered Scene

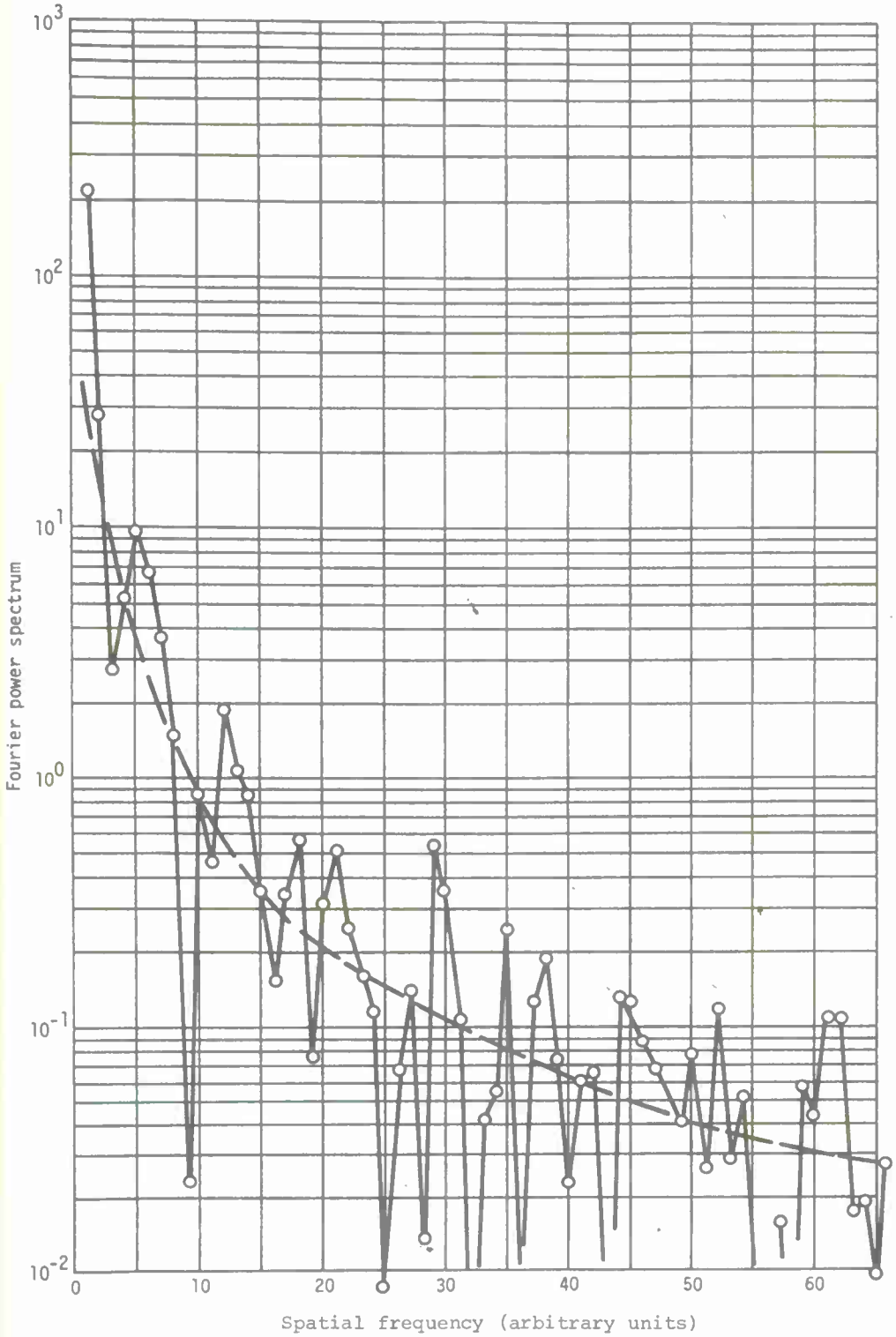



Figure 3. Power Spectrum Plot for a Typical Uncluttered Scene

## Scene Statistics

All models surveyed to date which included contrast thresholds also use a deterministic apparent or displayed contrast. It is argued in Reference 3 that the expressions involving contrast should be weighed by some form of statistics. Targets have, in a given scene, approximately constant emissivities and temperature distribution. Thus, the target/background contrast at the scene really follows the scene statistics of the thermal background. Similarly, at the display, the target/background contrast statistics roughly follows the brightness statistics of the scene. Reference 3 has compiled typical scene statistics as determined from filmed (display) imagery of the AN/AAQ-5 FLIR. This kind of statistics should be used to form a weighting function for the displayed contrast of a FLIR system.

## Sampling-Magnification Relations

Williams[23] has performed computer simulations of the effect of FLIR-like sampling on pattern recognition. Using a variety of subjects, he has determined that (almost independent of the number of target pattern samples) there is a system magnification which gives a local optimum pattern recognition performance. The optimum is "local" because, naturally, the absolute performance increases (in his tests) monotonically with sampling number.

These effects should be included in the FLIR target acquisition model, because the system magnification represents the interplay among sensor FOV, display size and viewing distance.

## Detection and Identification Sampling Requirements

Rugari[24] has reported FLIR airborne target acquisition measurement. These tests can be roughly characterized as taking place in a clutter free scene and in a high signal to noise situation. Thus the tests approximately test the pattern recognition performance of a man-FLIR combination. Rugari has found that the results can be modelled by invoking, as a stylized target, a rectangle which circumscribes the target. The pattern recognition requirement then is related to a minimal sampling of this area or of the geometric mean of its sides. Analyses have not yet been performed to see whether this measure of sampling requirements is in agreement with the bar target equivalent and/or the resolvable spots measures.

## FLIR Target Acquisition Measurement and Analyses

EOS and WPAFB have designed and conducted [14,25] an extensive and rigorous set of FLIR target acquisition measurements, using the Navy's Patuxent River E-O Test Facility. The tests included measurements of bar breakout and truck detection and recognition ranges. Backgrounds were varied and atmospheric temperature and humidity determinations were made. To maximize the experimental samples in each set of the test program, flight profiles were made essentially constant.

EOS has performed a variable analysis on the flight test measurements;[25] further analyses are in progress.

An alternate approach to the data analysis was presented by Mendez.[26] The approach is based on the assumption that $P_4$ (see Table III) represents the effect of SNRD on target acquisition and the hypothesis that the SNRD is the dominant factor in target acquisition. Then the MAFLIR flight test results are reduced to clear atmosphere for interpretation according to the following procedure:

1. $P_4(r_{meas}) = P_4(r_{calc})$, where

   $r_{meas}$=measured real atmosphere detection or recognition range,

2. $P_4(r) = f\left(K\frac{\tau_a(r)\cdot\Delta T}{MRT(r)}\right)$, where

   K=const.; $\tau_a(r)$=atmosphere transmittance at range r

   $\Delta T$=thermal signature

   $r_{calc}$=calculated clear atmosphere detection or recognition range

MRT(r)=resolution-sensitivity relation expressed as a function of range rather than spatial frequency.

3. For a given target and background, 1.) and 2.) reduce to

$$\frac{\tau_a(r_{meas})}{MRT(r_{meas})} = \frac{1}{MRT(r_{calc})} .$$

In this approach, it is seen that the MRT is interpreted as a transfer curve which can be used to reduce the flight test measurements to clear atmosphere. Then the reduced data can be interpreted in the sense of Johnson criteria (for example). The analysis procedure is depicted graphically in Figure 4. To estimate $\tau_a$ (r) from the temperature and relative humidity measurement in the MAFLIR program, a model atmosphere due to Barhydt[27] was used.

The approach described above makes it essential to fix and to know the MRT or RT of the system (preferably measured during the flight test).

The MAFLIR flight tests are difficult to interpret in a rigorous sense because the MRT of the system was altered in the course of the flight tests, and the experimental design made the RT one of the variables in the course of the flight test program. (Hopefully, future FLIR target acquisition measurement programs will exercise care in this regard). Nonetheless, analyses performed to date on these data indicate that the results can be interpreted in the sense of the bar target equivalent.

The algorithm described in this section is a first order approximation because it ignores angular subtense dependent threshold relations. Corrections of this type can be added by including a $P_3$-like term (see Table III) in the algorithm. This has not yet been done. More controlled measurements of the MAFLIR type are required.

## CONCLUSIONS

A survey of FLIR target acquisition models has been presented. These range in approach from analytical to phenomenological. Models of the "rule of thumb" type were not included.[28]

Most models are synthesized from analyses and intepretations of independent experiments performed to isolate the effect of a specific set of physical observables. The models do not exhibit agreement with respect to the interpretations or influences of pattern recognition and S/N (video or display).

The status and rigor of the modelling is dependent on test and analyses programs of the kind represented by the MAFLIR related efforts. Much more data of the MAFLIR kind, with better experimental design, is required in order to define an unequivocal FLIR target acquisition model which is valid from search through the higher tasks of target acquisition.

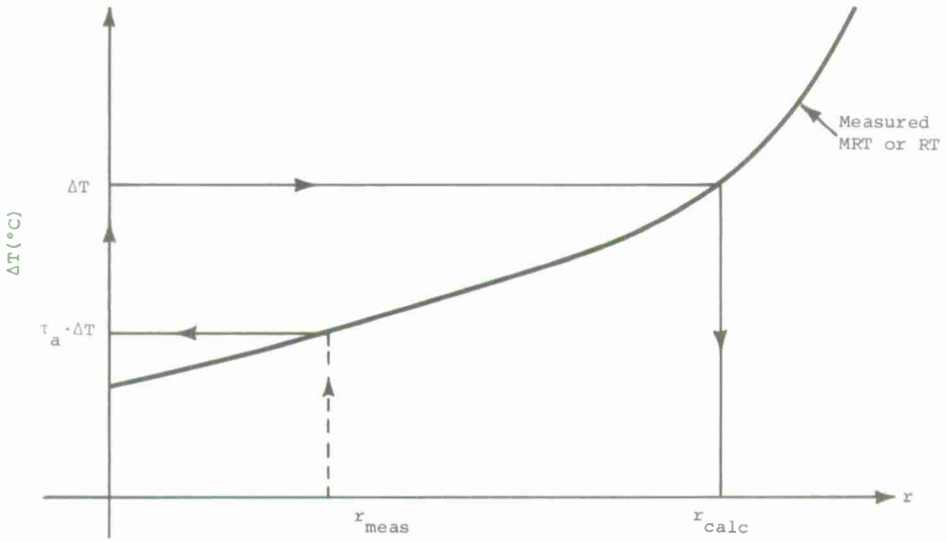



Figure 4. Graphical description of the use of the MRT (or RT) as a transfer curve to reduce real atmosphere target acquisition range measurements to clear atmosphere.

REFERENCES


1. Lloyd, J. M.,Thermal Imaging Systems. Unpublished manuscript, U.S. Army Electronics Command, Night Vision Laboratory. 10 May 1972.

2. Biberman, L. M.,A Guide for the Preparation of Specifications for Real-Time Thermal Imaging Systems. Institute for Defense Analyses Paper P-676, January 1971, IDA Log No. HQ70-11794.

3. Mendez, A. J.,Freitag, M. Z., Hallenback, R., Studies in Thermal Imagery (U). Martin Marietta Aerospace Tech. Report OR 12,125, Aug. 1972. (CONFIDENTIAL).

4. Bailey, H. H., Target Detection Through Visual Recognition: A Quantitative Model. The Rand Corp. - Memo. RM-6158-PR, Feb. 1970.

5. Schaefer, R. W., Aksamit, E. M., Heitzmann, R. K., et al. Multiple Airborne Reconnaissance Sensor Assessment Model (MARSAM II). Honeywell, Inc. Tech. Report ASD-RR-68-3, Part II, Feb. 1968.

6. Moser, Paul, Naval Air Development Center, Johnsville, Penn.; Private communication, 1972.

7. Pibble, W., IR Signatures - Scanning Radiometer, 3rd DOD Air-to-Ground IR Weapon Delivery Review, MICOM, Huntsville, Ala., 17-18 Oct. 1972.

8. Van Deman, G. L., and Wuster, W. L., A Mathematical Model for Estimating the Performance of Imaging Infrared Systems (U). IRIS, Vol. 15, No. 1, Sept. 1970, 257-269. (CONFIDENTIAL).

9. Pridgen, J. H., and Simpson, M. L., Interrelationships of Imaging Infrared Systems Performance with the Weapon Delivery Problem (U). IRIS, Vol. 16, No. 1, Oct. 1971, 481-490.

10. Isaacson, D. N., and Blackemore, J. W., A Mathematical Model Predicting the Probability of Detecting Targets Against Various Background Utilizing a Forward-Looking Infrared System (U). IRIS, Vol. 11, No. 1, Oct. 1966, 201-211.

11. Lapidus, B., Tactical Utility of Forward-Looking Infrared Systems (U). JDR, Summer 1970, 124-139, (CONFIDENTIAL).

12. Lawson, W. R., and Ratches, J. A., Thermal Imaging Systems Models (U). Paper presented at the Infrared Information Symposium, Imaging Specialty Group Meeting, Boston, Mass., Nov. 1972.

13. Bailey, H. H., Target Detection Through Visual Recognition: A Quantitive Model and Two Applications (U). JDRB, Spring 1971, 54-65. (CONFIDENTIAL)

14. Hodges, J. A., Infrared Image Test Program Study. AFAL-TR-70-269, AF Avionics Lab. AFSC, Wright Patterson AFB, Dayton, Ohio. (ED.)

15. Schade, O. H., A New System of Measuring and Specifying Image Definition, Proceedings of the NBS Optical Image Evaluation, Oct. 1951; NBS Circular 526, 1954, page 231.

16. Snyder, H. L., A Unitary Measure of Video System Image Quality. Paper presented at the Target Acquisition Symposium, Orlando, Fla., 14016, 1972.

17. Johnson, J., Analysis of Image Forming Systems. Proceedings Image Intensifier Symposium, U.S. Army Engineers R&D Lab., Ft. Belvior, Va., Oct. 1958.

18. Rosell, F. A., and Williams, R. H., Signal-to-Noise Ratio Thresholds for Image Detection, Recognition, and Identification. Paper presented at the Target Acquisition Symposium, Orlando, Fla., 15-16 Nov. 1972.

19. Rosell, F. A., Resolution-Sensitivity Model for FLIR. Proceedings Winter Meeting of the IRIS Specialty Group on IR Imaging. Sept. 1971.

20. Sendall, R., The Prediction of Signal Strength Required for Image Detection/Recognition on a Raster Generated Display. Paper presented at the Target Acquisition Symposium held at Orlando, Fla., 15-16, Nov. 1972.

21. Keesee, R. L., Detectability Thresholds for Line-Scan Displays. Paper presented at the Target Acquisition Symposium, Orlando, Fla., 14-16, Nov. 1972.

22. Saxon, J., and Higginbotham, H., Forward-Looking Infrared (FLIR) Scanner and Display Development. Tech. Report AFAL-TR-358, Feb. 1970.

23. Williams, L. G., and Graf, C. P., The Effect of Image Quality on Target Recognition. Paper presented at the Target Acquisition Symposium Orlando, Fla., 14-16, Nov., 1972.

24. Rugari, T., Airborne Target Acquisition Tests. Paper Presented at the Infrared Information Symposium, Imaging Specialty Group Meeting, Boston, Mass., Nov. 1972.

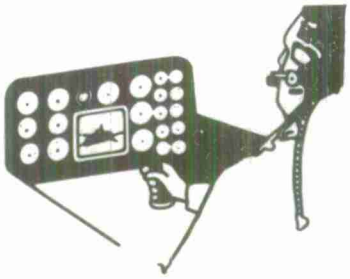
25. Hodges, J. A., Infrared Image Test Program - Variable Analysis. Air Force Avionics Lab. Tech. Report AFAL-TR-71-362, WPAFB, Dayton, Ohio, Feb. 1972.

26. Mendez, A. J., An Analysis of the MAFLIR Flight Tests. Paper presented at the Infrared Information Symposium, Imaging Specialty Group Meeting, NBS, Gaithersburg, Md., May, 1972.

27. Barhydt, H., Hughes Aircraft Company, 1971, Private Communication.

28. Lloyd, J. M., A Catalogue of Thermal Imagery Defects. Abstract submitted to the Target Acquisition Symposium, Orlando, Florida, 14-16, Nov. 1972.

29. Stathacopoulos, A. D., et al. Surveillance Systems Study: (Real-Time Night Reconnaissance/ Strike) (U). General Research Corporation, Report CR-437-2, 1967 (SECRET).



# VISUAL SEARCH FOR SYMBOLICALLY-CODED TARGETS

Warren H. Teichner
New Mexico State University
Las Cruces, New Mexico

and

Marjorie J. Krebs
Honeywell, Inc.
St. Paul, Minnesota

OFFICE OF NAVAL RESEARCH
TARGET ACQUISITION SYMPOSIUM
NAVAL TRAINING CENTER, ORLANDO, FLORIDA / 14,15,16 NOVEMBER 1972

# VISUAL SEARCH FOR SYMBOLICALLY-CODED TARGETS[1]

Warren H. Teichner
New Mexico State University
Las Cruces, New Mexico

and

Marjorie J. Krebs
Honeywell, Inc.
St. Paul, Minnesota

## ABSTRACT

Results of studies from the published literature were used as data for an analysis of tasks involving search for simple targets. Absolute values were plotted across studies on a series of common graphs. The relationships between the dependent measure (search time) and various independent variables were examined for systematic trends. The analysis yielded a number of findings including estimates of maximum scanning rate and information processing rates. Since these data were based on search for simple, well-learned target symbols, they are suggested as providing an estimate of the lower limits of search performance which can be expected.

## INTRODUCTION

There appear to be three kinds of visual search situations that have been studied experimentally. One is characterized by temporal uncertainty. The searcher knows where the target will appear, but he does not know when. Generally, when it does appear, the target is of short duration and easily missed if the locus of search at that moment is elsewhere. The appropriate measure of performance in this case is the frequency or proportion of targets detected. Research concerned with this problem has concentrated on the effects of target characteristics as they influence detectability, e.g. contrast, size etc., and with the question of decrements in target acquisition as the search period is prolonged (vigilance). We have recently considered aspects of both elsewhere (Teichner and Krebs, 1972, In press; Teichner, 1972).

A second kind of visual searching is characterized by positional uncertainty. In this situation the target is known to be present and it usually remains so until detected. What is unknown is the the location of the target. The appropriate measure of performance is search time, i.e., the time required to find the target. Of course, targets may have both positional and temporal uncertainty and these more complex situations comprise the third kind of search situation that has been studied. All three experimental paradigms are also representative of realistic situations.

In studies involving temporal uncertainty, i.e., Type I and Type III, both search time as well as the proportion of targets detected have been used as a dependent measure. It is important to distinguish between search time as used in those studies

and search time as used in the Type II case which involves only positional uncertainty. When targets are missed, search time reflects only those searches which were successful. Instances of missed targets contribute to the calculated proportion of detections, but not to the average search time. Thus, comparisons of search times may be, and usually are, not possible under conditions of equal target acquisition. On the other hand, in the Type II case where all targets are detected, from one comparison condition to another search time represents the same quality of target acquisition performance.

It would seem that somewhat different processes are implied by the two kinds of search time. Search time obtained from a study having only temporal uncertainty, at least in part, involves a process which leads to decisions of whether or not searching should be continued, whereas search time obtained from situations having only positional uncertainty would appear to be more purely representative of the search process as such. Both factors, those which determine the continuance of searching and those which determine the nature of searching, are critical to an understanding of target acquisition in all three situations.

This paper is concerned with search time in situations having only positional uncertainty. The experimental question, therefore, is that of how long it takes to find a target and not that of whether the target will be found. In the last few years there has been a considerable increase of interest in the problem beyond the context of target acquisition, and with respect to the nature of human information processing. Experimental investigations have been directed toward such issues as serial vs. parallel data handling, feature testing models, scanning models, and toward the nature of attention and short-term memory. Methodologically, the experiments require searching for a target or datum in an array of non-target stimuli. They differ from the conventional target acquisition study in that they de-emphasize properties of the stimulus and emphasize characteristics of the searcher.

Positional uncertainty studies may be viewed in still another context, the choice reaction time. More specifically, they may be seen as special cases of that type of choice reaction study in which the individual is required to respond to selected stimuli and required not to respond to any others. At this time we wish only to emphasize as strongly as possible that no matter how it is conceived, visual target acquisition involves both information processing and visual factors. Our effort was aimed, at the information processing aspect of the problem. Specifically, we were concerned with the question of whether available results are sufficient to provide relevant quantitative search time functions for Type II, the positional uncertainty case.

## METHOD

As with others of our recent studies, it was our hope to employ the absolute measures available in the literature as our data. By manipulation of those data we wished to determine if there is enough consistency within the literature to permit fitting functions which hold across the data of different studies. The data analysis starts with the assumption that only one independent variable determines most of the variation of the dependent measure. A failure to support that assumption forces the further assumption of two independent variables as primary. How far one goes in this direction depends upon what is available in the literature and how ingenious one is in manipulating it. It also depends upon success in establishing a literature-derived data base which contains the relevant results of quality investigations. Finally, what is relevant is determined by the restrictions of the investigation. For this study, the following constraints were imposed upon the selection of the experimental studies to be used:

1. The stimuli used had to be of an intensity such that they were easily detectible and

identifiable as individual stimulus elements.

2. The stimuli had to be simple and familiar elements such as letters, colors, and geometric shapes which might be expected to be known to subjects without specific learning. For alphanumeric stimuli only studies using single digits or letters were employed.

3. Studies involving masking, distorting, or any other form of noise were not used. However, studies using irrelevant or non-target, simple stimuli were used if those stimuli were part of the searched array.

4. Subjects had to be reasonably well-practiced at the task. In general, this was determined by the studies themselves since the amount of practice given prior to obtaining the reported data was highly variable. In two cases (to be discussed) the data were taken from the middle portion of experiments having unusally extensive practice in order to keep the learning factor approximately constant.

5. The amount of error of all kinds (omissions and commissions) had to be no more than five per cent. This criterion interacted with and partly determined what was considered to be "reasonable" practice. One minor exception was made in this regard and that, too, will be considered later.

6. From the subject's point of view the possible locations of the target had to be well-defined and known. This constraint precluded studies of empty fields.

7. All elements of the display, targets and non-targets, had to be present from the onset of viewing to the occurence of a response indicating detection.

## RESULTS

The reported data of most studies were in the form of the total time required to find the target. However, studies using more than one target per display reported the time to detect all of the displayed targets. As it turned out, we were unable to handle those studies in the same context as the others. Furthermore, those were too few, with too much interstudy variability, to warrant a separate analysis. Therefore, studies displaying more than one target at a time were discarded.

Some studies displaying only one target, reported the average search time per subgroup of stimuli in an array. For example, Neisser (1963) and subsequent studies based upon his findings (e.g. Neisser, Novick and Lazar, 1963; Cohen and Pew, 1970) reported the time per row in a matrix of alphanumeric symbols. Their data were converted to total search time, i.e. the time required from the start of a trial until the target was reported and to the time per stimulus or single symbol of the display.

It was also considered necessary to modify the description of the number of stimuli per array reported by all studies. All authors reported the average search time over a set of displays in which the position of the target was experimentally balanced within the set. For example, an experiment using an array containing 50 stimuli (49 non-targets and one target) would have balanced the position of the target over the set of displays. The mean search time of the set, then, represents a mean position location. The mean search time in this example is the search time for one-half of the number of elements in the array, i.e. 25. Therefore, the data to be presented which deal with the number of stimuli in the array will describe this variable as the "half number of stimuli."

With the restrictions listed above, a systematic search of the abstracted literature between 1960 and

1971, plus references from bibliographies, reviews and individual papers yielded 70 studies of which 12 provided 14 experiments having acceptable data. All of those studies required visual scanning of an array which was at a distance from the subject. Paper-pencil, letter-cancelling tasks were not accepted because the search time measures were assumed to be inflated by measurement with a stop-watch and by the manual aspect of the task.

We started by assuming that the only important variable was the number of stimulus elements in the array. As will be seen we also considered the size of the source from which the targets were selected (number of possible different targets), the size of the visual field, the nature of the stimulus elements, and the amount of stimulus information.

Figure 1 presents the total search time required to find a single target when only one kind of target is possible and when that target appears only once in the array. The results of eight different studies reporting 11 experiments are shown on a semi-log plot for which the abscissa is one-half of the number of stimuli actually present in the array. The four lines, drawn by eye, are intended only to provide a clearer picture of the trends. The study by Lehtiö (1970) differs from all of the other studies. In that study all stimuli including the target were digits and all varied in color and size. When the target was a particular color, it was the only instance of that color. When it was a digit it was the numeral 3, and the only 3 in the array, etc. Although the remaining studies in the figure are similar in that they are all cases of the target's differing or being selected from other similar stimuli on the basis of a single attribute, that attribute was the only difference between the target and the non-targets. Lehtiö's targets differed on two other dimensions as well.

Figure 1 represents a variety of specific kinds of target and non-target stimuli. Ignoring the results of Lehtiö, the figure suggests that the shortest search times were associated with finding a target circle which was larger than the remaining circles of the display. The next most rapid detection occurred when the target was a red circle among green circles of the same size for arrays having half numbers of stimuli greater than 100. The next longer search time was associated with finding a triangle in an array of circles, and the slowest search with finding a square among an array of circles.

The difference between the trends for a square target and a triangular target when each is embedded in an array of circles may be due importantly to the differences between investigative methods. In spite of that possibility, it is worth noting that a square is more similar to a circle than is a triangle and, therefore, that the square represents a target which should be harder to discriminate from the rest of the array.

The curve drawn through the triangular target data of Brody, Corbin and Volkmann (1960) and Volkman et al (1964) also fit the search times obtained from arrays using alphanumeric data. The one large exception is the study by Cohen and Pew (1970). That study may be compared directly with the one by Neisser, Novick and Lazar (1963) since it was an exact replicate in all respects except that the subjects were instructed to emphasize accuracy in finding the target whereas Neisser, Novick and Lazar's subjects and those of the other studies shown were given speed as the instructional emphasis.

The search time measures were usually reported in milliseconds. However, in order to reduce variability, Figure 1 and subsequent uses of search time as a dependent measure plots the data to the nearest .01 second.

In general Figure 1 suggests that at first search time is almost unaffected as the number of stimuli in the array increases, but that after some critical number, search time begins to increase. That observation was also made by Brody, Corbin and Volkman (1960). The inflection point

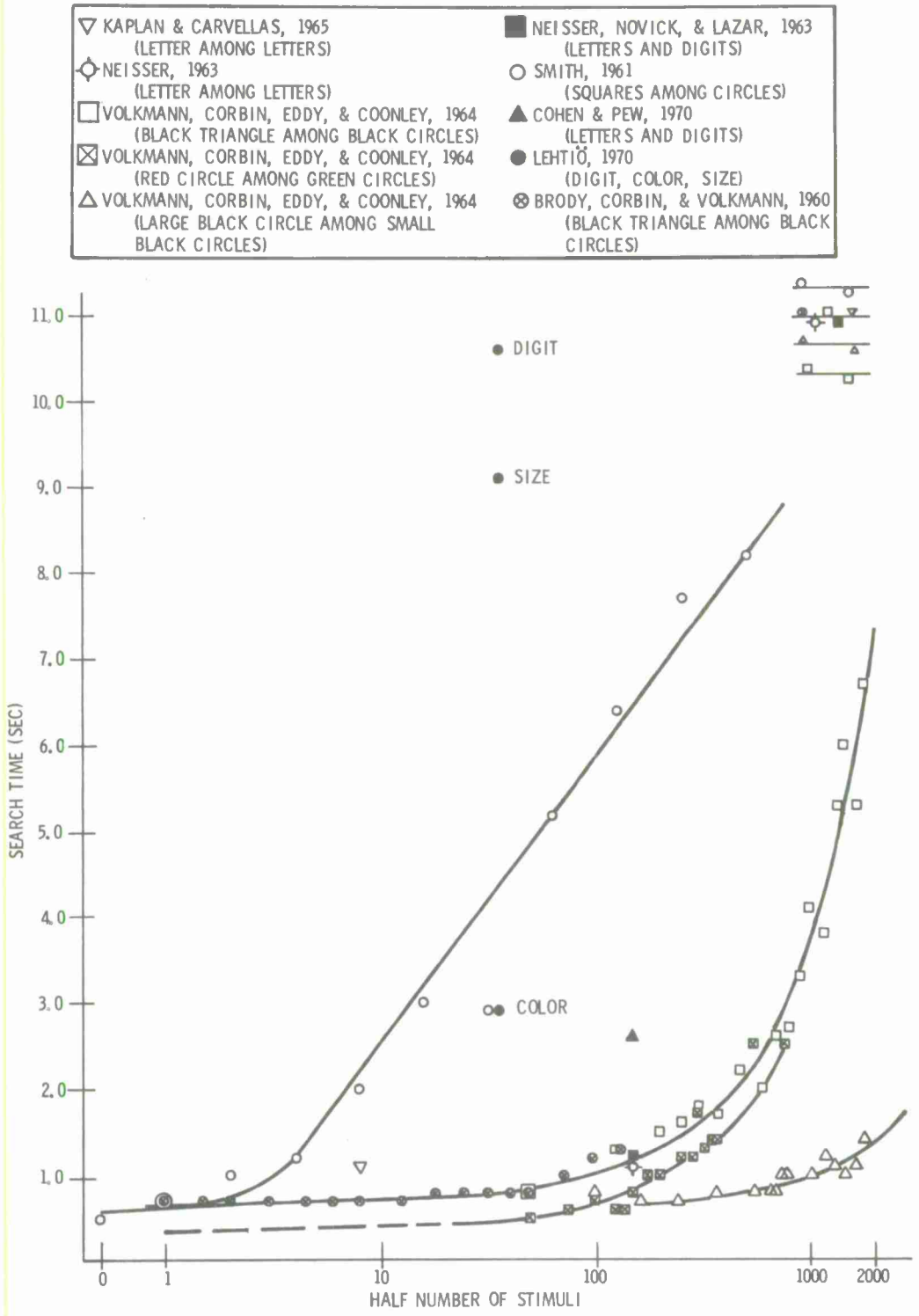



Figure 1. Search Time as a Function of the Half Number of Stimuli in the Display

and the slope appear to be related to the ease with which the target can be discriminated from its background. The more difficult the discrimination, the smaller the number of stimuli at the inflection point, and the steeper the rise after the inflection point.

The data of Lehtiö are completely out-of-line with any of the trends. His data are particularly interesting, however, since they suggest a different kind of search variable. That is, in all of the studies shown, excepting Lehtiö's, the subject could arrive at a decision by finding a stimulus which was different from the rest in any respect. Since in Lehtiö's study the stimulus elements differed with regard to combinations of two attributes, the subject had to find the target in terms of its unique attribute. After that, he may or may not have also acted selectively within the target. If we give credence to the difference in Lehtiö's results, it is reasonable to infer that his subjects were required to search for a specific attribute difference in an array of differences and, therefore, to inspect each stimulus as an individual source of information whereas in all other studies shown, the search was for any dishomogeneity in the array. The subjects of those other studies, therefore, could have swept the array in a fast scan. Results with simultaneous targets, reported by Eriksen (1962, 1953) provide additional support for this interpretation of Lehtiö's data. Thus, two different scanning strategies are suggested. One requires element by element inspection; the other allows for sweeping across the array. If they are thought of as opposite ends of a continuum, then variations of strategy might be in the number of elements to be inspected as a unit.

A second dependent measure of interest is the scanning rate, or its inverse value, time per element of the display, where an element is one symbol or stimuulus Time per element was calculated for the studies shown in Figure 1 and plotted in Figure 2. The figure shows very clearly that on a log-log plot this dependent measure decreases linearly as the number of stimuli in the array increases. The generality of the dependent measure is limited, of course, since the arithmetic involved would produce a decreasing measure even if the search time were constant over the whole range of the abscissa. That was not the case, although Figure 1 suggested that the search time may be constant up to a critical number of stimuli.

Figure 2 indicates that the observer appears to be capable of processing discrete elements at a rate of less than .0007 second per element. If rates of that order are within the individual's processing rate, then Figure 2 may be interpreted as conceivably representing an increase in the speed of serial processing. On the other hand, to the extent that the values are unrealistic, Figure 2 presents an artifact due to the arithmetic involved in calculating the dependent measure. Ignoring this question for the moment, and assuming that Figure 2 is meaningful, the figure indicates that as the number of stimuli in the display increases, the time spent per element decreases to a limiting value after which it remains constant. The limiting value appears to depend upon the nature of the stimuli used. In fact essentially the same limit appears to have been reached for the two middle trends of the figure. Inspection of Figure 1 shows that search time increased with a positive acceleration so that both trends are already almost asymptotic to an infinite search time. Looking at both Figure 1 and Figure 2 for those studies, it seems clear that the increasing search time is actually due to an increasing number of stimuli to process at a constant time per stimulus.

Figure 2 can be approached with more than one hypothesis. It can be supposed that after some critical number of stimuli, the observer shifts from serial to parallel processing. The two extreme trends at the high end of the abscissa would then suggest that for those tasks the shifting point had not been reached. The flat portions of those results could be interpreted as what happens when the shift occurs. The same change, unfortunately, is equally plausible as what happens when there is a shift in the

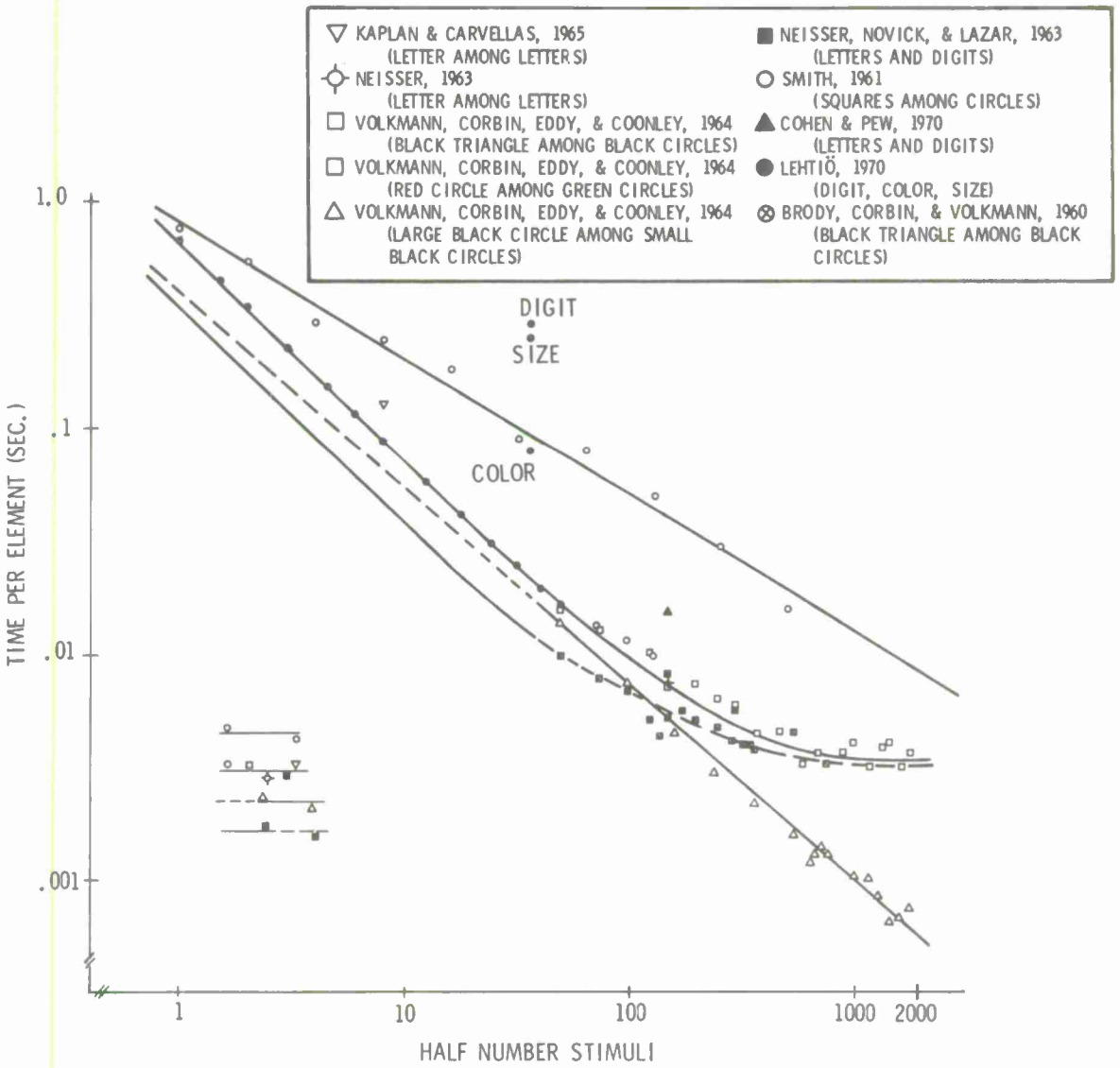



Figure 2. Time per Displayed Stimulus Element (symbol) as a Function of the Half Number of Stimuli in the display.

grouping or coding of multiple elements in the display. That is, as Vol,mann et al (1964) have suggested, the individual inspects each element of the display until a critical display number is reached after which the stimuli are scanned in groupings of some sort.

Volkmann et al (1964) have reported that as the density of the stimuli increased in their experiments, the search time also increased. Density was calculated as the total number of stimuli displayed divided by the visual angle of the field in degrees. Our plots of search time as a function of density and time per element as a function of density produced graphs which were very similar to Figures 1 and 2. For that reason only the latter will be presented.

Figure 3 relates time per element to density. As noted the figure closely approximates Figure 2. One important discrepancy should be noted, however. Of the four studies which used alpha-numeric displays, three used identical densities and one was almost at that density. Whereas in Figure 2, three of those studies were close to the trends of the other data and one was not, all four are now away from those trends. Furthermore, those four studies may now be seen to have provided a wide range

of values of time per element. Thus, at least for the data being used, Figure 3 does not provide support for density as an independent variable. Instead it suggests that it may appear to be such a variable only because of a spurious correlation with the dependent measure via its own dependence on number of stimuli.

The data shown in the previous figures were based upon searching for a known target when there were no alternative target possibilities. Figure 4 presents search time as a function of the number of possible different targets that could appear when only one of those targets is in the display. The parameter is the half number of stimuli in the display. It may be seen that for a constant number of displayed stimuli, search time increases as the number of possible targets increases. However, the data are not very well-ordered in terms of the parameter, number of stimuli.

Figure 5 relates time per element to the number of possible, different targets for the same data. With regard to the parameter, although the ordering is far from completely consistent, it is more systematic than it was on the previous figure. It is also interesting that the largest arrays are now at a constant time per element, i.e. they are independent of the number of possible targets. The small, 8-element array, on the other hand, still shows an increasing function of the number of possible targets.

Figure 5 is a little misleading in that time per element is plotted to the nearest .01 second whereas both Cohen and Pew (1970) and Neisser, Novick and Lazar (1963) reported their data to the nearest .001 sec. At that greater level of precision, their data also show an increasing time per element as the number of possible targets increases. The data being shown from those two studies are from the fifth day of practice.

The data of Kaplan and Carvellas (1965) are based upon any two runs through the series in one experimental session. By the tenth day

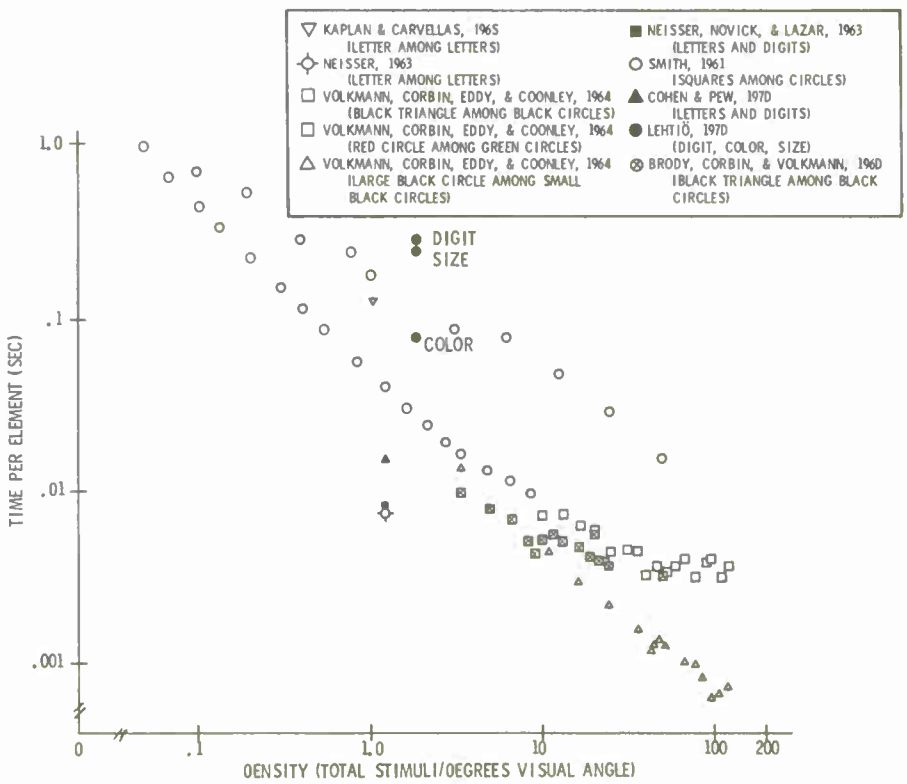



Figure 3. Time per Element as a Function of Stimulus Density

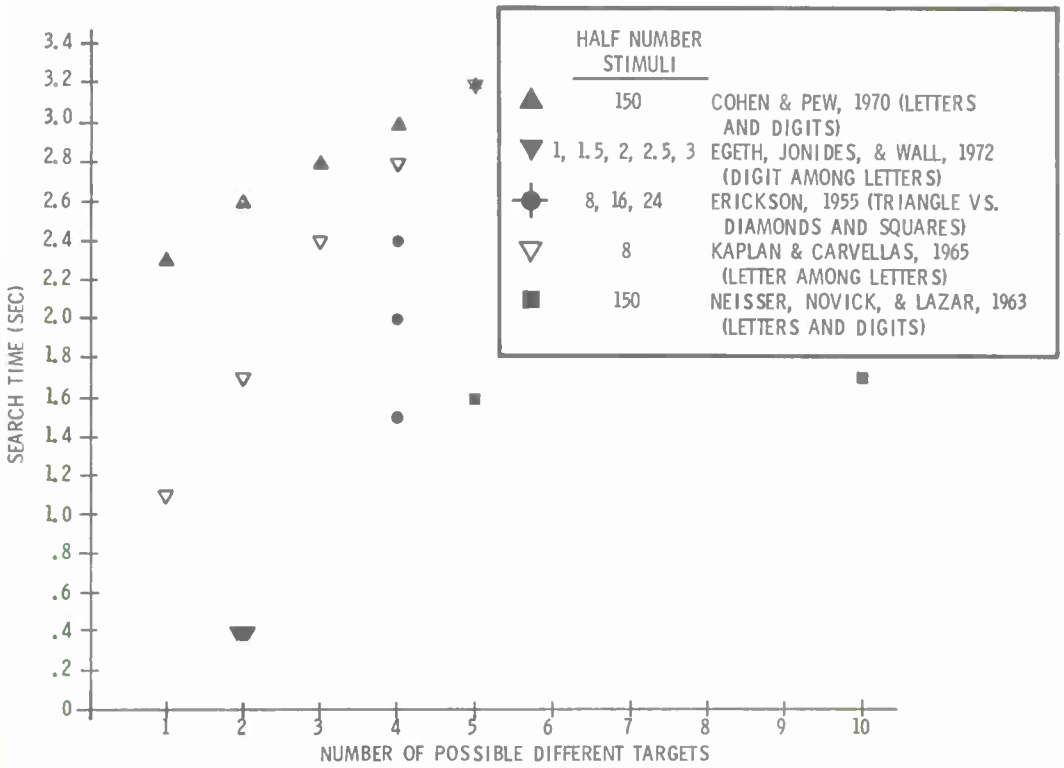



Figure 4. Search Time Related to the Number of Target Alternatives with the Half Number of Stimuli Displayed as the Parameter

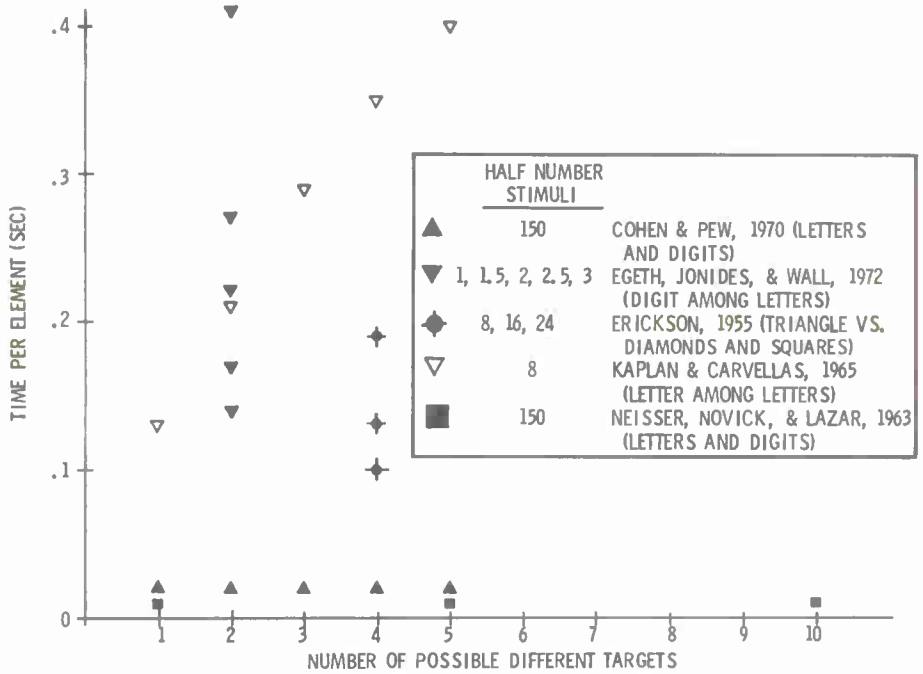



Figure 5. Time per Element Related to the Number of Target Alternatives with the Half Number of Stimuli Displayed as the Parameters

Neisser, Novick, and Lazar found that the time per element was the same regardless of the number of possible targets at least up to 10. Cohen and Pew did not find that result through 15 days of practice although with succeeding days the differences in time per element associated with the number of possible targets became markedly less. Presumably the differences between those two studies are due to the major experimental difference in procedure, i.e. instructions which emphasized speed (Neisser, Novick and Lazar) or which emphasized accuracy (Cohen and Pew). Regardless, the results suggest that with enough practice the effects of differing numbers of possible targets are lost. This suggestion finds very strong support in studies of choice reaction time which show the same phenomenon although only after much greater amounts of practice (e.g. Mowbray and Rhoades, 1959).

The visual search situation may also be analyzed in information theoretic terms and this has been done for the data of Green and Anderson (1956) by Meudell and Whiston (1970) with considerable success. Such an analysis may be based upon two different underlying search models. The first assumes that the individual is a random searcher. The presentation of elements of the display to the searcher then depends upon a random positioning of the eyes during scanning. In fact, data from studies of eye movements during search suggest that although the pattern of eye movements may be erratic it is far from random. A second factor involved in the particular studies that we are considering is that in all of the studies using alphanumeric symbols the subjects were instructed to search sequentially in the same manner as used for normal reading. Furthermore, it is a reasonable hypothesis that the eye movement sequencing in the other studies was at least strongly influenced by reading habits. On this basis a random search model seems less likely to represent what happens than a model which assumes a systematic search. In fact, that was the case in our analysis of these data and, therefore, only the results obtained with the systematic search model will be presented.

A systematic search model assumes that the subject inspects each element of the display in serial order. On inspection he makes a one bit decision, i.e. that the element is either in the target or in the non-target classification. If it is in the target classification, then he must make a further decision about which target of the possible alternative targets it is. In general, the amount of information, H(X), involved is:

$$H(X) = N + m \log_2 n \qquad (1)$$

where N is the number of possible locations or the number of symbols in the display if all possible locations are used.

m is the number of possible target alternatives

n is the number of elements or items in the source of targets and non-targets

As noted by Meudell and Whiston (1970) when the probabilities of the individual items in the source vary, the more accurate calculation of the amount information is:

$$H(X) = N + m \sum_{i=1}^{n} p_i \log_2 1/p_i \qquad (2)$$

where $p_i$ is the probability per source element. However, as they also note, and for the data involved here, the error involved in using Equation 1 when the probabilities are unequal is very small. Furthermore, for the present data it is not always possible to be sure of what the probabilities should be. For example, in some alphanumeric studies sources were established for experimental purposes which were selections of letters or letters and digits. Displays were then constructed from those sources. But, it is reasonable to assume that the subjects viewed those sources as if they contained all 26 letters or all letters and digits, and as if all source elements were equi-probable. The analysis to follow made that kind of subject-oriented assumption.

Figure 6 represents search time as a function of stimulus information

for all of the data of the previous figures and, in addition, for the results of three additional studies all of which involved more than one possible target.* Again, it may be seen that there is an increase in the dependent measure as the independent measure increases in magnitude and that more than one trend is evident. Also, although the calculation accounted for variations in the number of possible targets, the results of Cohen and Pew do not fall on any of the trends. Again, this is presumably because of the emphasis on accuracy which they employed. Except for the data of Lehtiö, as before, there is a general organization of the data into what appear to be three major trends. It is difficult to see what distinguishes among the trends, however.

*Egeth, Jonides and Wall (1972) presented an array of letters within which was placed a single digit as the target. The subject was instructed to report any digit as a target. Thus, the search was not for one of several possible digits, but of digit as a category. Accordingly, although the investigators selected from eight digit possibilities, we viewed the stimulus source as a one-bit source, i.e. letter or digit.

Figure 7 presents a plot of the data showing the relationship between time per element and amount of stimulus information. The measures are plotted to the nearest .01 sec. which accounts for the large number of points which are close to zero seconds. Again the data of Lehtiö are out of line with the rest of the data. Now, however, it can be seen that all of the other data, with the exception of those of Kaplan and Carvellas, fall on what appears to be a single trend. The

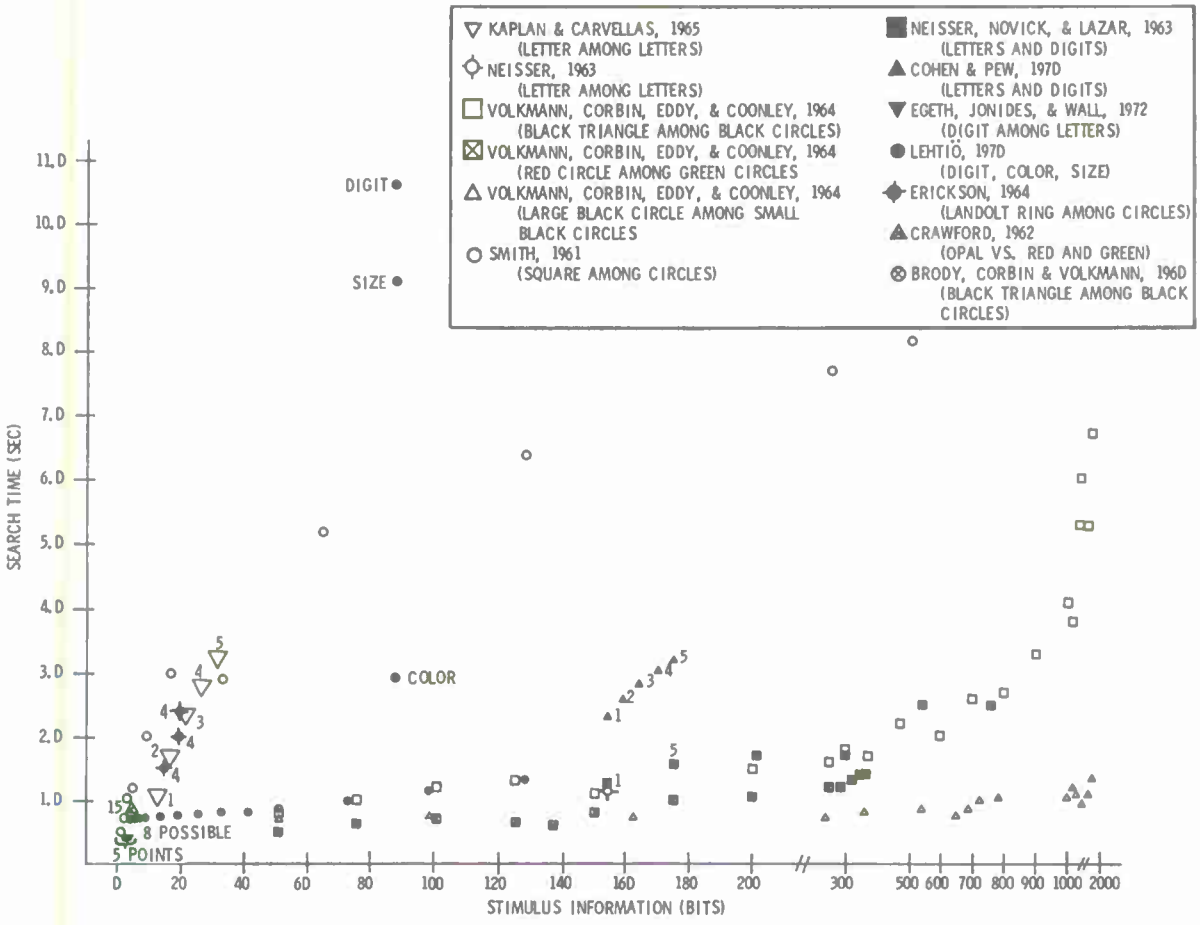



Figure 6. Search Time as a Function of Stimulus Information Content
Numerals Refer to the Number of Possible Different Targets

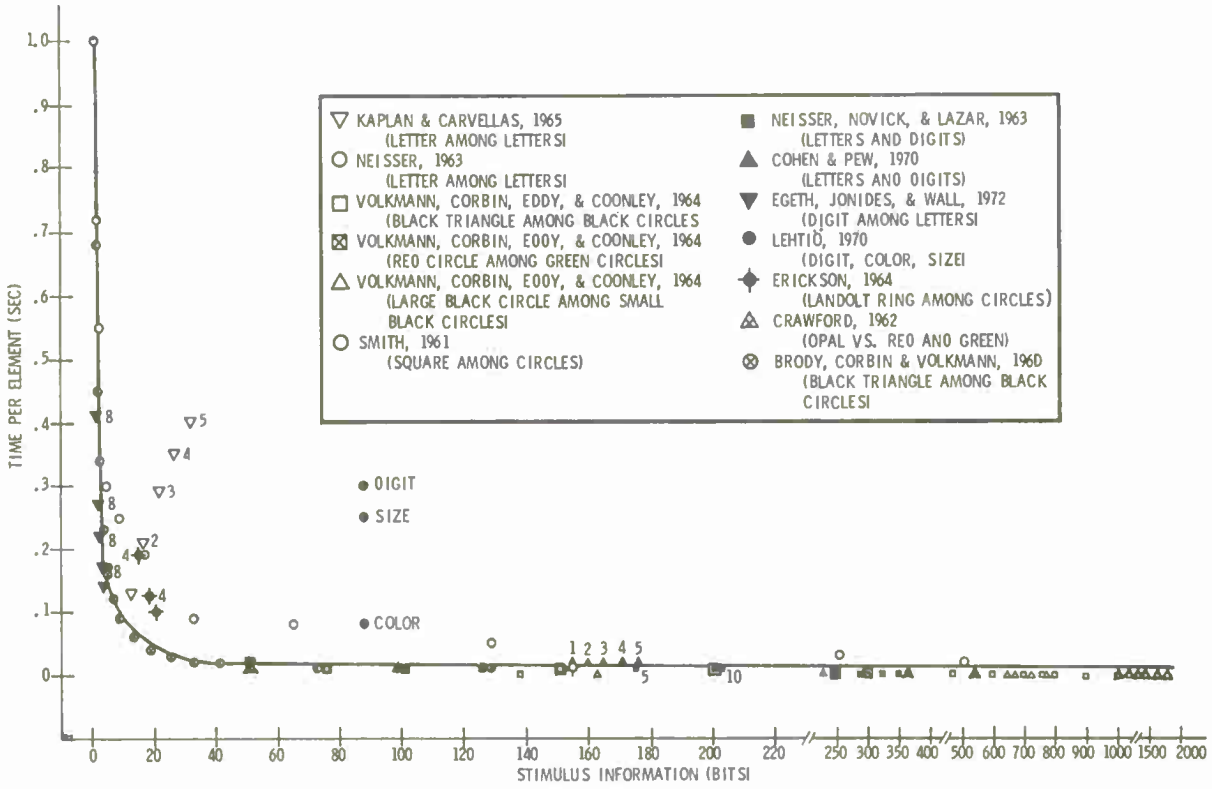



Figure 7. Time per Element as a Function of Stimulus Information Content
Numerals Refer to the Number of Possible Different Targets

smooth line, drawn by eye, ignores the results of Kaplan and Carvellas and of Lehtio. The function shows an extremely rapid decrease in time per element to 35 bits of stimulus information at which point the time per element is .02 seconds. The slight decrease in time per element after that is again, presumably, the artifact noted earlier in relation to Figure 3. This result suggests a maximum rate of 50 elements per second at a display complexity of 35 bits, assuming serial searching. It should also be noted that again the results of Lehtio are out of line, though closer. We have no explanation for the reversal in trend of the data of Kaplan and Carvellas.

Figure 8 presents the scanning rate as a function of stimulus information. The points were obtained by reading the smooth line of the previous figure and converting the values obtained to their reciprocals. The figure confirms the conclusion of the previous figure, but with additional emphasis. That is, the curve rises steeply to 50 elements per second at which point it reaches a limit. After that it is constant to about 55 bits where it begins to rise again. Had we been willing to plot time per element to the fourth decimal place, it would have been rising slowly, but continuously throughout. Thus, the rise beginning at 55 bits represents a change detectable at the second decimal place. But as we have suggested repeatedly, the mere addition of large numbers in the denominator of the quantity, search time/half number of stimuli = time per element, will produce continuously decreasing times per element even if search time is constant. Thus, we cannot accept the practice of reporting time per element to the third and fourth decimal places as the measure then becomes over sensitive to the artifact. Accordingly, we interpret Figure 8 as suggesting a maximum scanning rate or scanning capacity of 50 elements per second. It should also be noticed that the rate increases, the more the information to be processed up to that limit. Finally, in this connection, the curve appears to be discontinuous at the lower end,

i.e. the rates for two-and three-bit stimulus situations do not fall neatly on the remaining trend. We have omitted them from the smooth line, therefore. But whether this result should be viewed as error or as representing a different phenomenon is not suggested by the analysis.

The error rate of Neisser (1963) and of Neisser, Novick and Lazar (1963) exceeded our five percent criterion for the acceptance of experimental studies by a few percent. Assuming that the error rates of all of the studies are really very small of negligible, and by assuming that element information is what is being transmitted, the line of Figure 8 can be transformed to represent the rate of information transmission. The result of that transformation is shown in Figure 9.

The line of Figure 8 was transformed over the range from 4-35 bits of stimulus information and drawn as the solid line in Figure 9. The two unconnected points in the figure are transformations of the two- and three-bit values of Figure 8. The

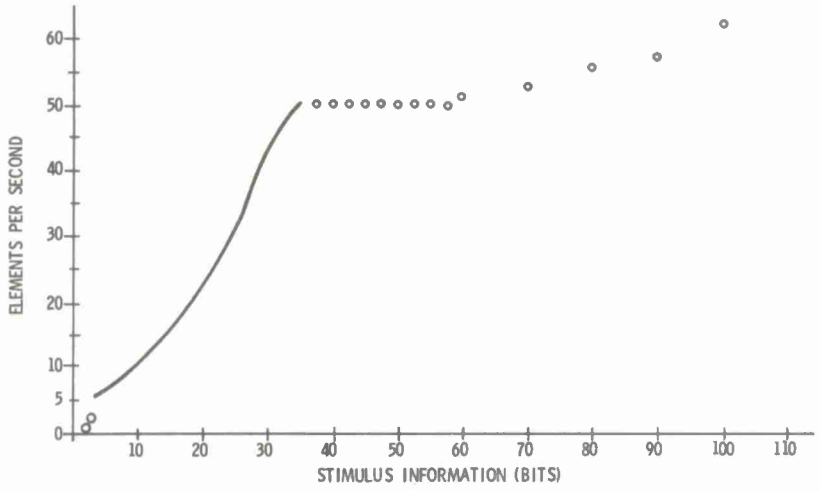



Figure 8. Scanning Rate as a Function of Stimulus Information Content

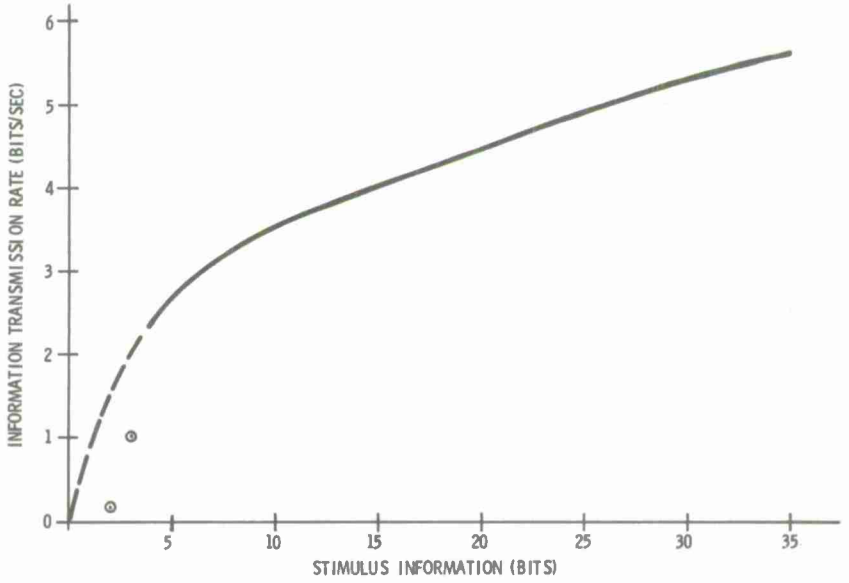



Figure 9. Information Processing Rate as a Function of Stimulus Information Content

dashed line is an extrapolation of the solid line to zero.

Again we cannot account for the deviation of the two and three-bit values from the main trend. In spite of that it is clear that the processing rate increases as the stimulus information content increases. The line indicates a maximum processing rate or capacity of 5.64 bits per second.

## DISCUSSION

In spite of the small number of experiments available which met our requirements, certain empirical relationships appear to have emerged from the data and some theoretical notions appear to be suggested. On the empirical side, the following conclusions appear to be warranted:

1. Search time for a target in an array increases, but the rate of searching also increases (and time per element decreases) as the number of stimuli in the array increases. The increase in search time is slow or even negligible at first and then increases rapidly. The inflection point and the rate of rise after it may depend upon the difficulty of discrimination between target and non-target elements. However, search rate varies directly with the number of stimuli in the display until a maximum rate is reached. The maximum rate depends upon the same factors as those which constrain search time.

2. Search time increases as the number of alternative possible targets increases. Up to some level of practice, the rate of searching also increases with the number of possible targets. After that level of practice is reached, the number of alternative targets, at least up to 10, probably makes no difference in search rate.

3. Search time is greater, when the emphasis is put by the searcher on accuracy of target selection.

4. For the kind of situations involved, density (number of stimuli per unit visual area) may be only spuriously related to search performance.

5. Time per element decreases as the amount of stimulus information in the display increases to a minimum time per element of about .02 second. Concurrently, scanning rate increases to a maximum of 50 elements per second.

6. The rate of information transmission increases with increasing stimulus information to a limit of 5.6 bits per second.

7. It is possible that the effects of stimulus information content on scanning rate and information transmission are quantitatively different below four bits of stimulus information from what it is above four bits.

On the theoretical side it seems apparent that a systematic serial search model may be appropriate to data of the kind used. Such data are from studies with no temporal uncertainty, with well-defined possible target locations, but with uncertainty about the location at which the target will be. The results of Meudell and Whiston (1970) also indicate that the sequential search model holds for more complex alphanumeric search tasks. The results also make it fairly clear that search time as a measure is less useful than scan rate since search time depends not only on the rate, but on the number of positions or symbols to be inspected.

It also appears that the conclusions suggested are restricted to cases where the target differs from non-target stimuli by a single stimulus attribute, and where only one target is present in the display.

Where attributes vary beyond that, scan rate and search time are considerably slower.

There has been a great deal of interest lately in the question of whether the search for multiple target possibilities is carried out in a serial or parallel fashion. Neisser, Novick and Lazar (1963) have suggested a serial process during learning of the task, but a parallel process with sufficient practice. The basis of their conclusion is their finding that with sufficient practice the rate of searching became independent of the number of target possibilities. If the data are used only to the nearest .01 second even the data of Cohen and Pew (1970) support that finding, and as we have already noted, such a finding is consistent with the effects of multiple alternatives in choice reaction time studies after enough practice. Whether or not parallel processing is involved is not suggested by our analysis. However, the analysis does suggest a scanning capacity of 50 elements per second for the conditions studied. The question now arises as to whether that rate is achieved because of the availability of parallel processing or whether a moderately high speed sequencer is involved. In any case, since the data represent the simplest cases, and therefore, since other search situations involving symbols, shapes, etc. will be slower, that rate is suggested as the maximum scanning rate to be used for practical purposes.

Although we really have only the one study of Lehtiö to go on, when those results are compared with the rest, there is the suggestion of a difference in searching strategy discussed earlier. That is, earlier we attempted to distinguish between a one-step, between-element search and a two-step, both between-and within-elements search. It might be supposed that which will be used will depend upon the degree of homogeneity of the stimuli. If so, the results do not suggest that homogeneity in a physical sense is what is involved although it could be an aiding factor. This appears to be the case since the physically heterogenous alphanumeric arrays yielded search times of the same order as one of the highly homogeneous arrays.

A more interesting hypothesis is of a psychological homogeneity which, for highly familiar stimuli, may be determined by the degree to which stimuli fall into the same psychological category. Accordingly searching for a letter among letters is more heterogeneous than searching for a digit among letters. It is interesting to note that in the study by Egeth, Jonides and Wall (1972), the target could be any digit. The display was one digit in an array of letters. That this study produced data which were among the shortest search times reported could be viewed as searching in a two-classification array for digits vs. letters, each letter being psychologically the same as any other letter, each digit being psychologically the same in its own category as target. The point is that once coded, digit vs. letters is psychologically no more heterogeneous than triangle vs. circles even though the difference in physical heterogeneity is very large. Our analysis indicates strongly that it is the psychological heterogeneity which is critical.

[1]Supported by a contract with the Office of Naval Research. We are grateful to Rae E. Brahlek of the American Institutes for Research who assisted in the early phases of literature analysis and to Nancy H. Martin who assisted in the data analysis. We are also grateful for the patience and encouragement of Dr. Martin Tolcott of ONR.

REFERENCES

Brody, H.A., Corbin, H.H. and Volkmann, J. Stimulus relations and methods of visual search. In Morris, A. and Horne, E.P. Visual Search Techniques. NAS-NRC Publication 712, Washington, D.C., 1960.

Cohen, V.V.R. and Pew, R.W. Small-size target sets in visual search under accuracy set. Midwest Psychological Association, 1970.

Crawford, A. The perception of light signals: The effect of the number of irrelevant lights. Ergonogics, 1962, 5, 417-428.

Egeth, H., Jonides, J. and Wall, S. Unlimited-capacity parallel processing of multi-element dis-

plays. Cognitive Psychology (In press) 1972, Research Contract N00014-67-A-0163-0001, 1971.

Erickson, R. A. Visual search performance in a moving structured field. Journal of the Optical Society of America, 1964, 54, 399-405.

Eriksen, C. W. Location of objects in a visual display as a function of the number of dimensions on which the objects differ. Journal of Experimental Psychology, 1952, 44, 56-60.

Eriksen, C. W. Object location in a complex perceptual field. Journal of Experimental Psychology, 1953, 45, 126-132.

Eriksen, C. W. Partitioning and saturation of visual displays and efficiency of visual search. Journal of Applied Psychology, 1955, 39, 73-77.

Green, B. F. and Anderson, L. K. Color coding in a visual search task. Journal of Experimental Psychology, 1956, 19-24.

Kaplan, I. T. and Carvellas, T. Scanning for multiple targets. Perceptual and Motor Skills, 1965, 21, 239-243.

Lehtio, P. K. The organization of component decisions in visual search. Acta Psychologica, 33, Attention and Performance III (A.F. Sanders, ed.), 1970, 93-105.

Meudell, P. R. and Whiston, T. G. An informational analysis of a visual search task. Perception and Psychophysics, 1970, 7, 212-214.

Mowbray, G. H. and Rhoades, M. V. On the reduction of choice reaction times with practice. Quarterly Journal of Experimental Psychology, 1959, 11, 16-23.

Neisser, U. Decision-time without reaction-time: Experiments in visual scanning. American Journal of Psychology, 1963, 76, 376-385.

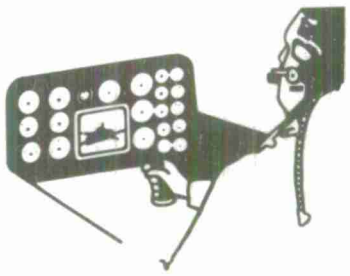
Neisser, U., Novick, R. and Lazar, R. Searching for ten targets simultaneously. Perceptual and Motor Skills, 1963, 17, 955-961.

Smith, S. W. Time required for target detection in complex abstract visual display. Memorandum of Project Michigan, Institute of Science and Technology, University of Michigan, 1961.

Teichner, W. H. Predicting human performance III: The detection of a simple visual signal as a function of time of watch. Technical Report 72-1, New Mexico State University, Las Cruces, New Mexico 88003, 1972.

Teichner, W. H. and Krebs, M. J. Estimating the detectability of target luminances. Human Factors, 1972, 14, 511-519. In press.

Volkmann, J., Corbin, H. H., Eddy, N. B. and Coonley, C. The range of visual search. Technical Documentary Report No. ESD-TDR-64-535, Decision Sciences Laboratory, Air Force Systems Command, United States Air Force, L. G. Hanscom Field, Bedford, Massachusetts, 1964, ASTIA Document No. AD 608 810.



# REAL-TIME TARGET ACQUISITION WITH MOVING AND STABILIZED IMAGE DISPLAYS

by

S. H. Levine and E. W. Youngling
McDonnell Douglas Corporation
St. Louis, Missiouri

OFFICE OF NAVAL RESEARCH
TARGET ACQUISITION SYMPOSIUM
NAVAL TRAINING CENTER, ORLANDO, FLORIDA / 14,15,16 NOVEMBER 1972

# REAL-TIME TARGET ACQUISITION WITH MOVING AND STABILIZED IMAGE DISPLAYS

BY

S. H. Levine and E. W. Youngling
McDonnell Douglas Corporation
St. Louis, Missouri

## ABSTRACT

Current high-performance aircraft achieve speeds which exceed man's ability to visually acquire a target before it is overflown. We are faced with the problem of extending man's capability to match that of the aircraft. Electronic sensors such as radar and TV systems offer potential solutions, however, considerable research into the best utilization of these systems is still required. McDonnell Aircraft's approach to this research has been to design and build a variable configuration display mockup (VCDM) in which actual flight images can be electro-optically manipulated to measure performance and evaluate system improvements. To maintain our operational orientation, stimulus imagery was generated using aerial imagery from the McDonnell Douglas Reconnaissance Laboratory data base, containing strategic and tactical targets.

## Moving Window Displays

In these systems, an area on the ground is imaged and moves down the display at a rate proportional to the speed of the aircraft. A series of studies was performed to evaluate the effects of image motion generated by various aircraft speeds. Imagery was viewed on the 5.5 inch CRT display in the cockpit station of the VCDM, at image motion rates yielding from 1 to 6 seconds on the display and simulated aircraft speeds of from 675 to 114 knots. Overall, the data indicates that a moving display can be used for the acquisition of "easy" targets (large truck parts, fortifications, storage area) at speeds resulting in as little as one second on the display. For all practical purposes, the relationship between image-motion and time-on-display limits the use of moving window displays to missions against large targets or those allowing slow aircraft speeds.

## Stabilized Image Displays

In these systems the sensor tracks an area on the ground so that a stationary image is displayed to the operator. As the aircraft closes with the target the image scale increases, producing a zoom effect. Since the system depends on tracking a fixed point, it can only be used for targets of known location, or where a specific ground area needs to be searched. The variables manipulated were: closing rate (360 to 1200 knots), starting field of view (100% vs 66% of original image), target offset from the center of the display (center, middle, edge) and type of briefing aid (photo or sketch of target). This series of studies indicated that targets can be successfully acquired on a stabilized image display at closing rates as high as 1200 knots, provided the navigation system is adequate to place the target in the center two-thirds of the field of view, and appropriate briefing aids are available.

In summary, at moderate speeds, moving window displays are effective against large targets, while stabilized image displays are effective against targets of known location at much higher speeds. This data contributes towards the definition of the aircraft envelope and the mission types within which man can effectively utilize these sensor/display systems.

## SCOPE OF PROBLEM

The human operator places severe limitations on the usable portion of the flight envelope of high performance aircraft engaged in air-to-ground target acquisition. These limitations are centered around difficulties in visual acquisition of ground targets. Laboratory studies indicate that a target size of from 10 to 20 minutes of arc is required for adequate (80% or better) target acquisition (Erickson, Main & Burge, 1967; Steedman & Baker, 1960; Boynton & Bush, 1955; Levine, Jauer & Kozlowski, 1969). Operationally this translates to an initial visual acquisition range of 3,300 to 10,000 ft. for a 20-30 ft. target. Operational and field studies obtained initial visual acquisition ranges of from 5,000 to 11,000 ft. for vehicles in the 20-30 ft. size range (Blackwell, Omart & Harcum, 1959; Dyer, 1964; Wade, 1964). Allowing for the difference in target acquisition procedure, altitude, and velocity, the laboratory and the operational results are in good agreement.

Relating these acquisition ranges to operational situations, an aircraft with a velocity of 300 knots will cover 5,000 ft. in approximately 10 seconds. Allowing for the best case acquisition range (11,000 ft.), at this velocity the aircrew will have only 22 seconds between target acquisition and target flyover. During this time the aircraft must be maneuvered into firing position, the target must be acquired on the weapons guidance system, lockon achieved and weapons launch initiated.

An unclassified summary of time line analyses of the critical weapons delivery functions for a number of mission configurations using a typical lockon before launch missiles showed that an average of 19 seconds was required for launch (Dolce, Whiteside & Wright, 1970). Using this data as a boundary condition, man's envelope was plotted against altitude and velocity to allow a direct comparison with the operational envelope of a typical high performance attack aircraft (Figure 1). This comparison clearly shows that the aircraft potential greatly exceeds man's capability. The designers of real-time and near-real-time reconnaissance and reconnaissance/strike systems are faced with the problem of extending man's target acquistion capability to approach the aircraft capability. The majority of the research on extending man's performance envelope has centered on real-time and near-real-time displays utilizing sensing aids such as radar and long-focal length electro-optical systems. These systems significantly extend man's target acquisition range and have the added benefits of recording an expanded energy spectrum and a capability for automated data processing. However, they also bring up a new series of problems relating to system utilization and the observer's ability to acquire targets from the displays system outputs. These problems can be summarized as the selection of sensor systems and the display of their outputs in a way which maximizes target acquisition and acquisition range and minimizes time to acquire and false alarms. Because of the large number of variables involved, this task is far more difficult than it would seem. Each element in the system can be analyzed and evaluated with respect to its own set of variables and to some extent the effects which these variables have on the system elements interfacing with it (Figure 2). For example, the target/background contrast on the ground and the sensor sensitivity will determine sensor output; the sensor output and the display capability will determine target/background contrast on the display; and in turn the target/background contrast on the display and the observer's visual capability will determine the perceived target/brightness contrast. The acquisition systems may be studied by relating variables from any system element with any of the others, however, in doing this all intermediate elements must also be evaluated. This approach results in complex analyses whose results are often equivocal and difficult to apply.

In the course of our research we have evolved three concepts to aid in the study of the acquisition problem and reduce some of the complexity.

First, the experiments should concentrate on the display image/observer interface (Figure 3). This interface - where the eyeball meets the display - has elements of both the hardware and observer elements of the system. The observer's response defines system performance in terms of targets acquired, and provides a direct measure of system goodness. The

display image represents the end result of the hardware part of the system. Once a display has been identified which produces adequate target acquisition it is possible to work backwards to select values of other hardware elements which produce the required display. This yields a simplified display image - performance oriented solution to the acquisition problem.

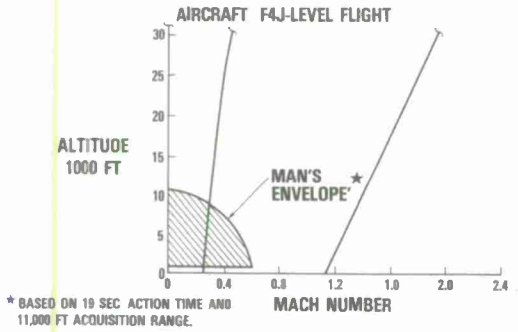



Figure 1. Man's Visual Capability Limits the System

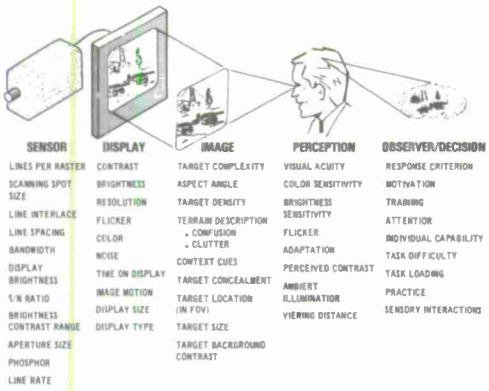



Figure 2. Target Acquisition Variables

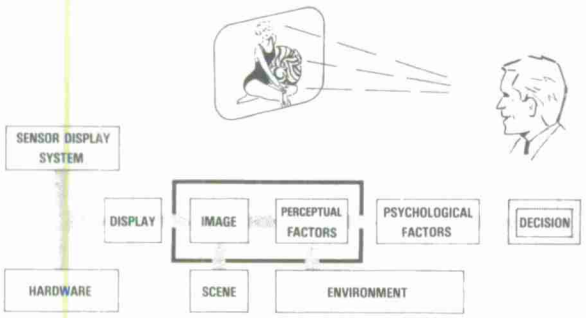



Figure 3. Concentrate on This Interface

Second, attempts should be made to identify summary metrics or compound variables which account for performance. These variables must be empirically measurable at the display face and directly relatable to performance. Our research has identified one such summary metric, resolvable lines-over-target (Figure 4) (Levine, Jauer & Kozlowski, 1970). This measure is derived from J. Johnson's (1958) observations concerning resolution requirements for target recognition/identification and consists simply of the product of the target height, in millimeters, as imaged on the screen, and the number of line elements per millimeter that can be resolved by the observer viewing the display under the same conditions as would occur during target acquisition. Analyzing target acquisition performance from 2 separate in-house experiments revealed a correlation of from .81 to .88 between performance and resolvable lines-over-target (Figure 5). This measure seems to offer a simple and effective estimate of expected performance, given a known display situation and target size. In concept, it is closely related to the MTFA which yields a predictive measure of overall performance with a variety of target sizes.

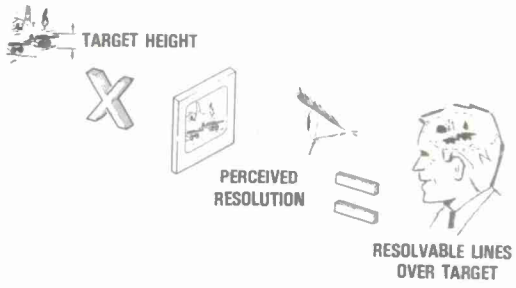



Figure 4. Summary Metric

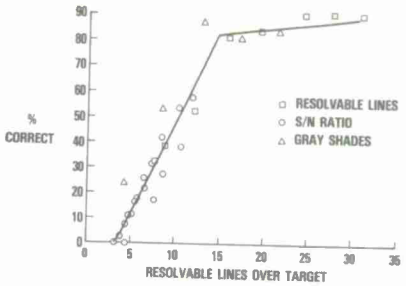



Figure 5. Summary Metric

MTFA is derived from the modulation transfer function (MTF) of the display and the contrast demand function of the human observer. The MTF is a plot of the resolving power of the display, with respect to sinusoidal resolution patterns, at various modulations or contrast levels. For CRT displays as contrast decreases the display can image progressively smaller resolution elements. Engineers have found it very useful because the MTFs of all the elements in a display system can be combined mathematically to obtain the MTF of the final display without actually putting the system together. The contrast demand function (CDF) is a measure of the contrast the observer requires to discriminate between resolution elements. For the observer, as contrast increases the ability to perceive progressively smaller resolution elements increases. Plotting the two curves against increasing modulation and resolution yields a monotonically decreasing function for MTF and a monotonically increasing function for CDF. The area encompassed by these two functions is the MTFA, and represents the resolutions that the display can image and the eye can see, at various contrast levels. This measure is used to predict the goodness of a display with respect to a population of targets, like those found in the real world. It is a predictor of average performance. The resolvable lines-over-target measure takes a single measure of the display/observer resolution capability at a single contrast and relates this to a specific target characteristic, target height. Both MTFA and resolvable-lines-over-target are based on the same perceptual principles, discriminable resolution, however, the former yields a general prediction of display-observer performance while the latter makes predictions concerning specific targets.

Third, the relationships between experimental findings relating to the display/observer interface should be directly related to aircraft/sensor system parameters (Figure 6). This relationship should be explicit and couched in language appropriate for design engineering. While the findings of specific studies are sometimes only general approximations of the true relationships, they are often the only data available on the interrelations between display system parameters and target acquisition performance. As such, they can play an important role in guiding the thinking of engineers responsible for hardware selection and design. Just as scientific data must be communicated to the scientific community to become "real," technological data must be communicated to the engineer to be useful.

| SYSTEM ELEMENT | VARIABLE | OBSERVER/DISPLAY |
|---|---|---|
| AIRCRAFT SPEED | RATE OF IMAGE MOTION | IMAGE MOTION<br>TIME ON DISPLAY |
| NAVIGATION ACCURACY | TARGET PLACEMENT IN SCENE | TARGET LOCATION ON DISPLA<br>TIME ON DISPLAY<br>MAXIMUM TARGET SIZE |
| SENSOR FIELD OF VIEW | SCALE AND GROUND AREA IMAGED | TARGET SIZE<br>LOCATION ON DISPLAY |

Figure 6. Explicate Display/System Observer Relationships

The research program which led to these recommendations was part of an internally funded series of studies on the display problems specific to long range real-time sensor systems. This evaluation has considered, to date, both moving and stabilized image applications of such systems. The following section reports on experiments concerning both of these image display systems.

## MOVING IMAGE DISPLAYS

The utilization of long focal length electronic sensors in conjunction with real-time and near-real-time displays will extend man's target acquisition range by functionally magnifying the scale of the terrain to be searched. This, in theory, should allow faster aircraft speeds while maintaining or increasing the time to target flyover. These systems are deployed aimed out of the aircraft at some fixed downward look angle. Using this method of deployment the image motion across the display is proportional to the aircraft speed and high speeds result in rapidly moving, short time on display images. In order to effectively evaluate the constraints this image motion places on the system, a series of studies of target acquisition and image motion were conducted. These studies considered target difficulty, target cues and target reference images, and their interrelationship with image motion.

Two exploratory experiments were conducted to determine target acquisition of military targets with natural backgrounds.

Targets typical of low and mid-intensity conflicts, such as bridges, truck parks, fortifications, and AAA sites, were selected from KA-80 high panoramic imagery of Southeast Asia. Twenty-five inch strips containing the targets were cut from each frame and spliced together to make a stimulus film. The film was advanced and converted to a video signal with a Flying Spot Scanner, built by the McDonnell Douglas Reconnaissance Laboratory. The output signal from the scanner was a standard 525 line, 2:1 interlace, 30 frame per second video output. The subjects viewed this image on a "9" Conrac TV monitor mounted in a simulated cockpit (Figure 7). Image presentation and motion were kept uniform through automatic control. The CRT image was 5.5 inches high with a scale of 1:2500. The image motion provided a total target time-on-display of 1, 2, 3, 4, 5, and 6 seconds which was equivalent to aircraft speeds of 675, 338, 227, 169, 135, and 113 knots, respectively. The performance measure was percent correct identification of the targets. Correct identification was achieved by the subject's correctly indicating that the target was on the display.

Figure 7. Cockpit Mockup and Display

In the first experiment 24 targets and 12 subjects were used. Based on target size, type, concealment and background clutter, targets were divided into difficult and easy groups. The difficult targets were revetments, trench fortifications and small truck parts and had a display size of approximately 1/5 inch. The easy targets were 3/4 inch on the display and consisted of large truck parts, forts, and fortified positions. Prior to the run, subjects were read the instructions and given a review of the target keys. Each target had a key with a print of the target and a brief written description. The keys covered only the target and the immediately adjacent ground area, and were available to the subjects during the entire experiment. The results are presented in Figure 8. Performance decreased as speed increased for both the hard and easy targets at motion rates greater than 1.7 inch per second. At slower rates, performance was nearly constant. The performance curves indicate that, for the display size and scale used, viewing times greater than three seconds did not significantly improve performance. The three second value appears to be independent of target difficulty as the curves for the easy and difficult targets are nearly identical in shape. While a major determiner of the absolute level of performance, target difficulty does not appear to interact with the other display variables. This would indicate the possibility of improving performance based on the target difficulty-image rate tradeoff curve. For example: a "difficult" target at a given scale might be viewed at twice the scale. This would decrease the viewing time to one-fourth the original, i.e., from 6.0 to 1.5 seconds. However, if the scale increase transferred the target category from "difficult" to "easy," the performance would be increased from 55% to 80%. Additional studies varying image scale are needed to fully evaluate this possibility.


Figure 8. Study One Moving Window Display

In the second experiment ten additional difficult targets were selected from imagery similar to that used in the first experiment. All ten targets were of the same approximate size and shape; however, five of them were situated near prominent terrain features such as roads or rivers.

The other five targets had no unique geographical features to aid in their acquisition. Twelve subjects were used. Prior to each run, the subjects were given a one minute period to study target briefing photos. Each set of briefing photos consisted of an image of the actual target area and images of two areas very similar to the target. Unlike the first experiment the briefing material was not available during the test run. Image motion rates of 675, 227, and 135 knots were used. All other factors were the same as in the first study. The results are graphically summarized in Figure 9. The data indicate that an increase in viewing time results in increased performance for the targets located near a prominent terrain feature. For the other targets, image motion rate had no effect on performance, and performance was generally below 10 percent. The prominent terrain feature apparently served as a context cue to the target's location. The improvement in performance for targets with natural cues lends credence to the concept of providing artificial cues to improve performance. Performance data for the ten most difficult targets from the first experiment is also plotted in Figure 9. This gives a comparison of performance when target keys are available during a mission. A comparison of the two curves indicates that the presence of target keys yields performance similar to that obtained when the target is near a major context one.

Analysis of these findings with respect to system utilization indicate that aircraft speeds yielding more than 3 seconds on display are required for good performance. The system does not appear to be adequate for small targets or targets of opportunity, although if context cues or target keys can be provided performance will improve.

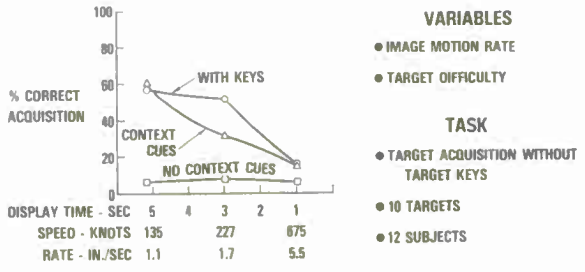



Figure 9. Study Two Moving Window Display

## STABILIZED IMAGE DISPLAYS

A narrow field of view (FOV) image stabilized electro-optical sensor has also been suggested as a means of extending man's visual performance envelope. This sensor will theoretically allow acquisition of targets at extreme standoff ranges, 10-20 miles. In operation, the sensor system will be turned on at extreme range and aimed through the use of the aircraft's navigational system. Theoretically the range of this system is limited by atmospheric effects and the relationship between aircraft altitude and the earth's curvature. Practically, however, the effective range of the system will be determined by the systems operator's ability to locate and recognize the target on the display. This ability is a complex function of a large number of variables derived from display, sensor, and observer centered parameters.

A number of studies were performed to determine the systems operator constraints on the utilization of the system, and where possible to develop techniques and procedures which will extend its operational utility. Our intial efforts were directed toward establishing a performance baseline and defining the boundary conditions for target location and acquisition. To implement this program, a simulation of a long focal length airborne TV system was configured with the McDonnell Reconnaissance Laboratory's Displays Mockup Facility. The sensor was simulated with a 70 mm Slide projector, rear projection screen, and a closed circuit television (CCTV) camera system with a remotely controlled zoom lens. The equipment set up is shown in Figure 10. The input scene was projected onto the rear projection screen and imaged by the zoom lens into the TV camera. By controlling the rate-of-zoom lens' setting, a simulation was made of a fixed focal length system in a moving aircraft at each of several speeds. The rate of zoom was controlled with a position servo system that tracked an input signal proportional to the zoom factor as a function of time. The optical axis of the lens was held fixed with respect to the projected scene, simulating the effect obtained with a ground stabilized sensor guidance system. The subjects viewed the display on a 9 inch, 525 line, Conrac TV monitor mounted in the cockpit mockup (see Figure 7).

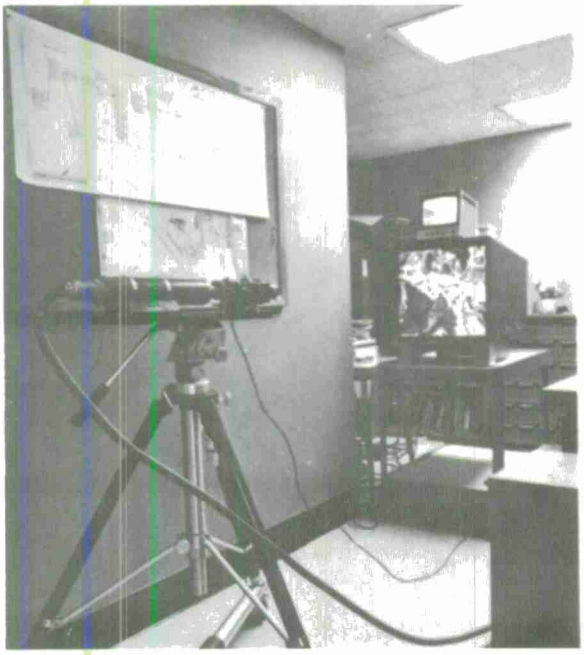


Figure 10. Stabilized Image System Simulation

Forward oblique imagery taken from the Reconnaissance Data Base was used to construct 70 mm (2-1/4 x 2-1/4 inch image) slides of 18 typical targets. Using the estimated resolution losses of the photo-copying processes, the optical projection system and an assumed minimal resolution of 20 lp/mm in the data base imagery, it was concluded that the limiting factor of the display image quality was the CCTV system, just as it would be for most real-world situations. The targets included buildings, POL storage tanks, water towers, and electric power substations. Image scales were chosen to produce a constant 1/4 inch target size on the cockpit display at the onset of each trial. With a constant displayed target size, we defined a baseline which can be used with the true target size, to calculate an acquisition range for other size targets.

The output of the zoom system was presented on a 5.5 inch square area marked off on the display monitor. At the start of a trial, the projected image, at minimum scale, to represent the sensor coverage at long range, was presented on the cockpit display and the observer began his search for the target. As the trial progressed, the CCTV system zoomed in on the center of the image with the rate of zoom appropriate to the simulated speed. This system simulated a narrow (approximately one to two degree) FOV sensor at 20 miles initial range and a 20,000 feet altitude. The targets were acquired by slewing a cursor to a position over the target image and pressing a fire button on the cursor control joystick. Target keys were available for reference at all times during the trial.

The controlled variables were aircraft speed, target location in the FOV, and type of key. Aircraft speeds of 360, 560, 900 and 1200 knots were simulated. When the target was in the center of the display, zooming in enlarged the displayed target. When the target was offset from the center, zooming in increased the offset as well as the displayed target size. For sufficiently large offsets the target went off the display (out of the sensor FOV) as the lens zooms in on the center of the image. Three concentric bands were used to identify target location in the FOV. The bands were from center to 1/3 the distance to the edge of the image, from 1/3 to 2/3 the distance to the edge of the image, and from 2/3 to the edge of the image. Each band contained six targets at random azimuths.

Target briefing cues, target keys, were also varied. One group of 12 subjects received a 15X enlargement of the target and immediate surrounding area printed on a 5 x 7 inch glossy photograph. Another group of 8 subjects received a line sketch of the target area traced from the 5 x 7 photo (Figure 11). The line sketch had greater than 100:1 data compression when compared with the photo. This is significant when transmitting images over an electronic data link system.


Figure 11. Picture and Sketch Target Acquisition Aids

Time to acquire the target and target acquisition accuracy were measured. The data were plotted as a function of aircraft speed with the effects of offset and target keys as parameters and an analysis of variance was performed. The data indicate that accurate target acquisition can be obtained if the target is in the center 2/3 of the FOV (Figure 12). Time to acquire the target did not vary with offset from the center of the image. All offset positions showed an approximately linear drop in acquisition time of about 15 seconds as speed varied from 360 to 1200 knots. This lack of position effect on time-to-acquire the target is not surprising as the targets were originally screened by offset groups to select six targets having a minimum spread of acquisition times. Photo keys were superior to line sketched keys in both acquisition time and percent of correct acquisitions (Figure 13). Overall a 10 percent reduction in performance and a 7 second increase in time were found using the line sketch keys. An evaluation of the offset by briefing aids interaction, however, indicated that while there were no significant time effects, there was a significant performance interaction. The effects of the sketch versus the photo briefing aid were not significant for the center 2/3 of the image, thus if the navigation accuracy can put the target in the center 2/3 of the image, the sketch aid can be utilized with little or no penalty in percent correct acquisition.

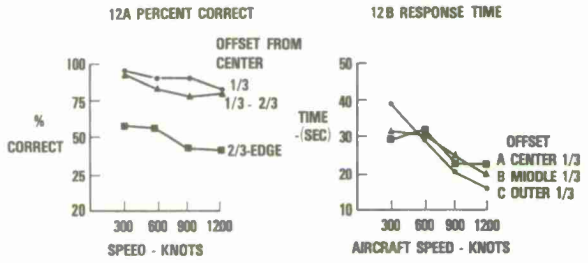



Figure 12. Study One Effects of Offset

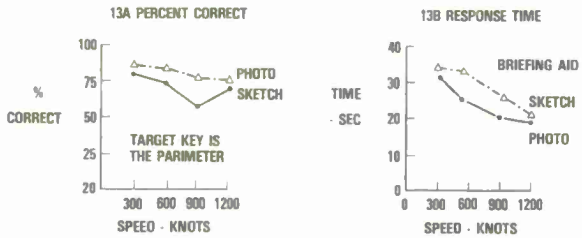



Figure 13. Study One Effects of Briefing Aids

The general effect of aircraft speed was a 15 second drop in acquisition time and a 10 percent drop in performance over the speed range of 360 to 1200 knots. The maintenance of a good degree of accuracy at the 1200 knot speed indicates a great potential for this sensor system in high performance aircraft.

A second study was conducted to evaluate the effects of changes in FOV on performance. This study used the same procedure and photo briefing aids as the first study. The variables manipulated were closing rate, navigation accuracy and FOV. Closing rates of 360, 560, 900 and 1200 knots were simulated by automatically presetting the zoom rate on the CCTV system. Fields of view of 1.30 and .90 were used, the latter resulting in an increase in the scale of the image and a decrease in the ground coverage. This was simulated by starting the zoom procedure at different magnifications. The change in starting point resulted in image viewing times of 45, 58, 98, and 158 seconds for the 1.3° FOV and 30, 39, 65, and 106 seconds for the .9° FOV. In all cases the time-on-display for the narrow FOV exceeded the average time for acquisition for the wide FOV, insuring that the narrow FOV did not excessively distort performance due to differences in presentation time.

Performance was measured with respect to time to acquire the target and target acquisition accuracy. This data was plotted with respect to the effects of FOV as a function of aircraft speed and appears in Figure 14.

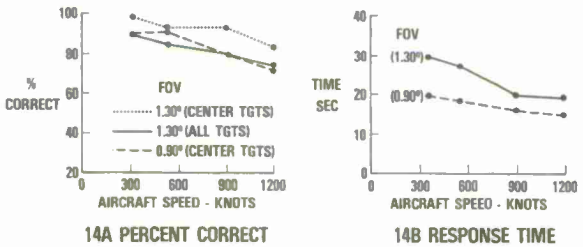



Figure 14. Study Two Effects of FOV and Closing Rate

The area covered by the narrow FOV contained only those targets in the center 2/3 of the wide FOV scene, however, it distributed their locations over the entire image. Comparisons were therefore made of both a truncated 12 target sample (those

targets in the center 2/3 of the image) and the complete 18 target sample (Figure 14A). The number correct for those targets which appeared in both FOVs varied as a function FOV, with the wide 1.3° FOV yielding better performance. For this case, the targets all appeared in the central 2/3 of the display and as such had long display times and a favored search position, both of which typically increase performance. If the narrow FOV data is compared with the full compliment of wide FOV targets, so the target distribution over the image is the same (Figure 14A), the overall difference between the two FOVs is not significant. An analysis of the time to acquire the target indicates that faster acquisition times are found with the narrow FOV (Figure 14B). This is most probably due to the increased scale associated with this FOV.

Overall, relating the findings of both studies to system design, a sensor pointing accuracy sufficient to place the target in the center 2/3 of the display will give the best acquisition performance. Changes in FOV can be accounted for by changes in placement in the scene and increases in scale. Since this is a zoom system small scale targets will be seen at larger and larger scales as the aircraft closes with the target. This factor washes out the effects of initial target size, provided the target remains on the display as the system zooms in on it. Placement in the center 2/3 of the image ensures both adequate scale and search time during the simulated approaches. Additional work is needed to define precise limits of this relationship with respect to initial scale and other zoom rates.

Target acquisition from aircraft with speeds as high as 1200 knots are possible utilizing this system. However, at these higher speeds the available time approaches the threshold of required time for critical combinations of actual target size and sensor's angular resolution. Converting the findings into an operator envelope and comparing it with the visual envelope (Figure 15), we find a potential doubling of the usable portion of the aircraft envelope.

The line sketch target key offers a 100:1 data compression advantage at a cost of only 10 seconds time and 10% performance degradation. If the navigation accuracy requirements are met and the target is in the center 2/3 of the display the sketch can be used at no cost in performance degradation. The narrow bandwidth required to transmit the line sketch key gives this technique a special potential for use in a hunter killer team to destroy fleeting target.

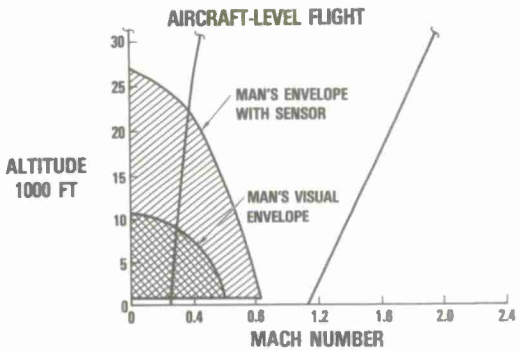



Figure 15. Sensors Can Double Usable Flight Envelope

REFERENCES


1. Blackwell, H. R., Ohmart, J. G. and Harcum, E. R., Field and Simulator Studies of Air-to-Ground Visibility Distances, Rept. 2643-3-F, University of Michigan Research Institute, Ann Arbor, Mich., 1959.
2. Boynton, R. M. and Bush, W. R., Laboratory Studies Pertaining to Visual Air Reconnaissance, Pt/1, WADC Tech Report 55-304, Wright Air Development Center, Dayton, Ohio, 1955.
3. Dolce, S. L., Whiteside, G. A., Wright, H., Multisensor Weapon Delivery Subsystem, AFAL-TR-69-314 Vol. 1, Air Force Avionics Lab., Wright Patterson AFB, Dayton, Ohio, 1970.
4. Dyer, G. C., Effects of Aircraft Speed on Low-Altitude Acquisition of Ground Targets (Phase II), APGC-TDR-64-40, Eglin AFB, Florida, 1964.
5. Erickson, R. A., Main, R. E. and Burge, C. G., Airborne Television Monitor Evaluation, Notes TP 4209, U.S. Naval Ordnance Test Station, China Lake, California, 1967.
6. Johnson, J., Vision Transforms and Elementary Decision Making in Sixth Annual Army Human Factors Engineering Conference, U.S. Army Engineers Research and Development Laboratory, Ft. Belvoir, Va., 1960.

7. Levine, S. H., Jauer, R. A., and Kozlowski, D. R., Human Factors Requirements for Electronic Displays: Effects of S/N Ratio and TV Lines-Over Target, McDonnell Douglas Rept. No. A0217, McDonnell Douglas Corp., St. Louis, Missouri, 12 Jan. 1970.
8. Levine, S. H., Jauer, R. A., and Kozlowski, D. R., Observer Performance with TV Imagery: Gray Shades and Resolution, McDonnell Douglas Report No. H 398, McDonnell Douglas Corp., St. Louis, Missouri, 30 Sept. 1969.
9. Steedman, W. C. and Baker, C. A., Target Size and Visual Recognition, Human Factors, 1060, 2, 120-127.
10. Wade, J. E., Effectiveness of Training Procedures for Low-Altitude Target Acquisition and Recognition, APGC TR-64-68, Eglin AFB, Florida, 1964.

# A SYSTEMATIC APPROACH FOR PREDICTION AND IMPROVEMENT OF TARGET ACQUISITION PERFORMANCE

By

J. D. Gilmour
The Boeing Company
Seattle, Washington

OFFICE OF NAVAL RESEARCH
TARGET ACQUISITION SYMPOSIUM
NAVAL TRAINING CENTER, ORLANDO, FLORIDA / 14,15,16 NOVEMBER 1972

# A SYSTEMATIC APPROACH FOR PREDICTION AND IMPROVEMENT OF TARGET ACQUISITION PERFORMANCE

J. D. Gilmour
The Boeing Company
Seattle, Washington

Abstract: A representative example is cited of the coordinated use of complementary research methods to systematically improve quantitative prediction and operational performance in target acquisition. A modeling concept is described that provides for the organized treatment of problem parameters, an evolutionary data management system, and the identification of leveraged parameters offering significant improvement in prediction and/or performance. Illustrative applications of analysis, simulation, laboratory testing, and field test validation are then used as examples of how prediction capabilities and system performance can be improved relative to the parametric implications of existing models. The examples used are scaling and classification of target backgrounds, and head-up target predesignation using onboard navigation system information.

## INTRODUCTION

The performance of current operational systems is severely restricted by a limited ability to acquire and identify surface targets in time to permit the effective delivery of ordnance. The visual search requirements imposed on the pilot, who must collect and process large quantities of visual and/or sensor-mediated information in real time, typically result in target acquisition performance levels substantially below those of which the system is theoretically capable. Technological solutions to this problem require an improved parametric understanding of the operational, environmental, design, and crew performance factors affecting system performance, the identification of high-leveraged system variables offering significant improvement potential, and the systematic evaluation of performance improvement concepts under real-time man-in-the-loop conditions characteristic of actual mission situations.

Progress towards improved prediction and performance in the target acquisition task can be greatly facilitated by the coordinated use of several complementary research methods. Parametric modeling serves as a useful tool for organizing problem variables and qualified data. It also focuses attention on quantitative data deficiencies, and tends to indicate the areas of research that potentially could lead to significant performance improvements. Techniques of analysis, laboratory testing, and simulation can then be applied as appropriate, singly or in combination, to satisfy high-leveraged data requirements or to test a given improvement concept. Finally, controlled field tests can be used to verify and/or calibrate selected points in the overall performance matrix. It is probable that each of these methods, when combined in a systematic and coordinated fashion, contribute more to an improved knowledge of target acquisition than any single method used independently. This is especially true where an economical expenditure of resources is a prime constraint.

## A MODELING FRAMEWORK

While a detailed model description is beyond the scope of this paper, a general explanation of the modeling scheme in use will assist in showing the inter-relationship between it and other

research and analysis efforts. This modeling scheme was initially developed for application to the visual target acquisition problem (Gilmour and Emerson, 1965), and has since been used as a general format for detailed parametric modeling of both the day-visual and night/all-weather tasks. Although the basic framework remains the same, models dealing with radar, TV/LLLTV, and IR/FLIR additionally interpose the appropriate sensor/processor/display variables between the operator and his environment. Only the generalized format will be considered here.

To be widely useful, it is desirable that an acquisition model organize the problem parameters in relation to operational performance criteria that have a real world counterpart. In this way, the objectives of the model can be oriented toward defining the functional relationships between operational, environmental, design, and aircrew variables and the mission performance of the system. The ultimate criteria in this case is acquisition performance, but several intermediate criteria can also be established which provide for the systematic and logical treatment of parameters and which give recognition to the highly sequential nature of the acquisition task.

In the present case, the operational performance criteria of the model were defined in terms of a listing of sequential requirements for successful target acquisition performance. These requirements included four major elements: (1) Geometric Intervisibility - providing geometric line-of-sight and compatibility between available field-of-view and the ground area containing the target, (2) Visual Target Availability - providing a supraliminal presentation of the target element to the observer, (3) Real-Time Search - successful visual sampling of the target element from all of the dynamic scene elements present, and (4) Discrimination and Decision - requiring detail discrimination to a level necessary to satisfy the observer's pre-stored "target" definition.

It can be noted that these requirements are cumulative and sequentially-dependent with respect to successful acquisition of an assigned target. It is also evident that certain groups of variables will be operative in determining whether each particular sequential requirement is met. (However, categories of variables are not necessarily exclusive to a single requirement, since some variables can be seen to operate on different requirements in more than one way.) This permits the operational performance criteria of the acquisition model to be specified in terms of the particular group of contributing parameters. A general schematic of the resulting model is illustrated in Figure 1, showing three intermediate performance criteria as exposure, detection and identification submodels. (These terms are used largely for convenience in notation; the reader is invited to substitute other labels if desired.) The exposure and detection submodels correspond to the first two requirements listed above, while the latter two requirements have been combined in the identification submodel. Each submodel addresses a specific question relative to the sequential requirements for successful acquisition. Each also treats a specific group of variables, examples of which are listed in Figure 1.

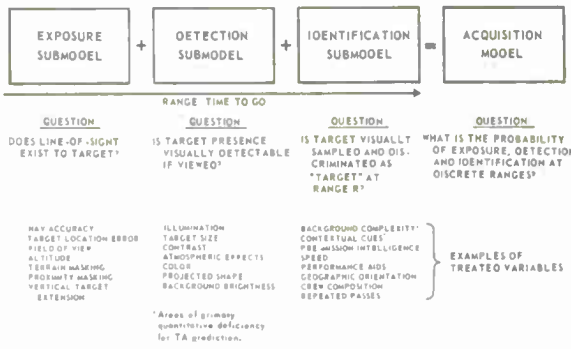



Figure 1. Target Acquisition Prediction Model

The sequential nature of the acquisition task is further illustrated by placing the above model on a range/time continuum. This is shown in Figure 2, where target range and time-to-go decrease progressively to the right. In the hypothetical example shown, the upper limits of acquisition performance are constrained by target availability (the combination of exposure and detaction criteria). Actual performance (identification), however, characteristically occurs much later than this upper limit.

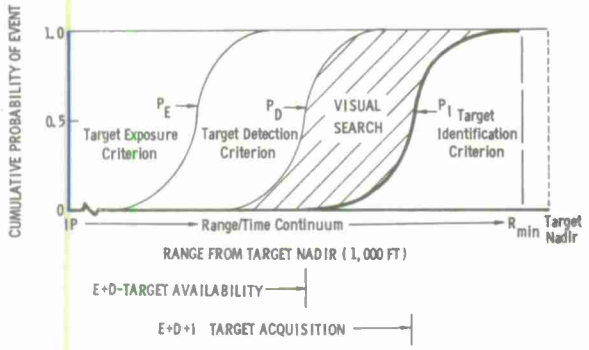



Figure 2. Acquisition Model Range/Time Sequence

Prediction of the visual search and discrimination performance that accounts for this difference has never been particularly successful, largely due to our inability to deal quantitatively with the variables operating in the identification submodel. This has been particularly unfortunate both because the unpredicted performance variations in this region are typically very substantial, and also because a large number of delayed or unsuccessful acquisitions occur that cannot be systematically explained or corrected. It may, on occasion, be worthwhile to consider the upper limits of performance attainable in a system, but where the system seldom if ever performs this way operationally, it is a mistake to take this as a routine expectation.

## TARGET/BACKGROUND SCALING AND CLASSIFICATION

One parameter contributing to the identification submodel analysis has been particularly troublesome. This factor has to do with the embeddedness (or, conversely, conspicuity) of a particular target in a particular background or setting. Both simulation and field test data consistently show extremely wide performance variations that are associated with so-called target differences. Such variations may often comprise 70-80% of the total test variability. An example of this is shown in Figure 3, taken from the simulator validation test portion of a large-scale field test.

The performance variations represented are, of course, not strictly a matter of differences between the target items themselves. Rather, they represent the total problem of visual search in a dynamic and complex visual environment, and hence include factors related to targets, backgrounds, and the interactions between the two. The central question in visual search is the time (and range) required by the observer to visually sample the desired target item from the barrage of visual information available. This will be determined largely by the characteristics of the background or target setting (terrain, cultural development, contextual cues, etc.) and by the target's relationship to these characteristics.

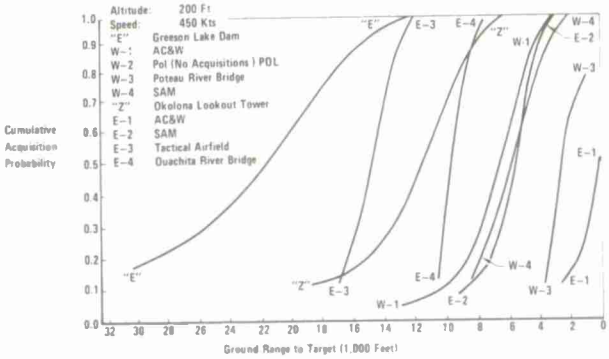



Figure 3. Cumulative Acquisition Probability as a Function of Range for 10 Prebriefed Targets

This is illustrated in the acquisition performance data shown in Figure 4. Performance is seen to differ widely on two similar "bridge" targets located in different settings. Cultural development and terrain patterns differed considerably in the two cases.

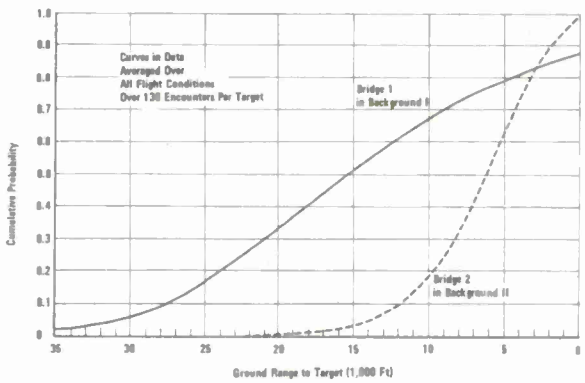



Figure 4. Acquisition Curves for Two Bridges - Different Backgrounds

As a further example of the effects of visual search requirements on acquisition performance, Figure 5 shows performance in a typical target/background encounter plotted as a percentage of maximum available range of the target. The use of approximately 60% of the available range/time for search, while reasonably characteristic, varies substantially under different target/background conditions.

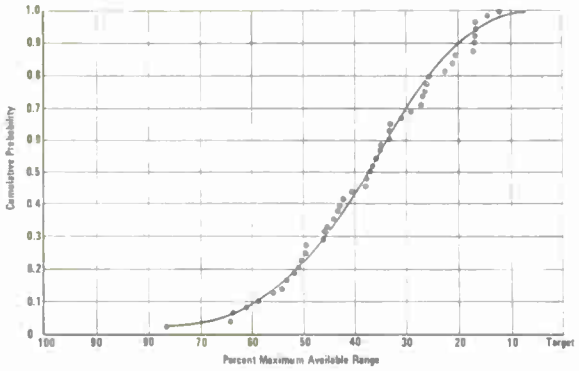



Figure 5. Cumulative Probability Curve Fit to a Set of JTF-2 Test Data (Target W-4)

In a study recently completed for the Office of Naval Research (Zaitzeff, 1971), initial work was begun on a technique to predict target/background performance effects. This study made use of 160° 70mm color motion picture and reconnaissance imagery from JTF-2 target acquisition field test courses, and of dynamic simulator acquisition performance data which had previously been validated against large-scale military field test results (Wyman, Gilmour, and Snyder, 1968). For purposes of this study, static scenes were extracted from the cinematic imagery at 10 discrete ranges relative to each of 10 targets. Acquisition performance data consisted of encounters of each of the 10 targets by a total of 160 operationally-ready military aircrews.

The purpose of this study was to begin development of a target/background scaling and classification metric that could be used to predict visual search performance over a wide range of target/background conditions. The initial objective was a preliminary algorithm relating measured characteristics of target/background scenes to the actual visual search performance of aircrews when acquiring these targets under real-time conditions.

The approach followed was to investigate a number of measurable properties of each of the scenes representing different stages of the target encounters. Analyses were then performed to evaluate the independence, quantitative effect, and relative predicting power of these variables relative to the criterion performance. Finally, a regression equation was developed that offered significant prediction capabilities, economy, and reasonable prospects for stability when generalized to new target/background situations.

The 15 candidate classification parameters investigated are listed in Figure 6 according to the measurement techniques used. They are largely self-explanatory except for the subjective scaling measures. Separate studies using 16 subjects each were conducted to evaluate the number of potential confusion (target-like) elements in each progressive scene (ambiguity), the percentage of observers designating the target with a high confidence level for a given scene (static acquisition), and the degree to which individual scenes were judged to be visually nonuniform on a ten-point scale (heterogeniety). In the last instance no target was specified; in the other studies pretest briefing materials were limited to the target item and its immediate surround.


Figure 6. Candidate Target/Background Classification Parameters

Analyses performed to compare measured target/background characteristics with the criterion dynamic acquisition performance are summarized in Figure 7. Factor analysis was used primarily as an investigative tool to examine the correlations among the original set of independent variables and the criterion. Moderately high correlations were

observed in most cases, and were generally in the expected direction. A rotated factor matrix, however, revealed a total of 7 orthogonal factors with a distribution of loadings that suggested a moderate amount of interdependence among the set of variables and the existence of a number of redundancies. This in turn suggested that almost as much predictive power could be obtained from a considerably reduced set of variables.

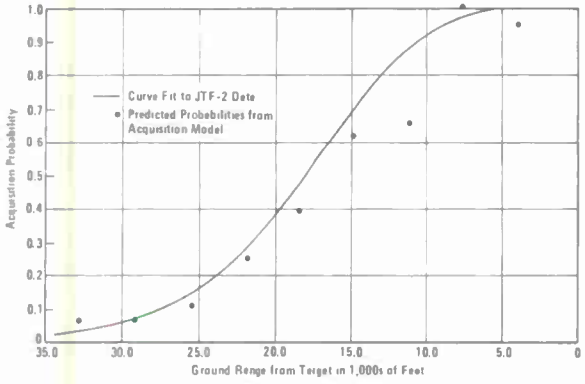



Figure 7. Target/Background Analyses

Linear stepwise regression was used to determine the quantitative contributions of the original variables to the prediction of acquisition performance on the targets used in this study. A regression equation using all 15 variables was seen to account for 93% of the criterion variance, while an equation using a reduced set of 7 of these variables was still able to account for 90% of this variance. Due to the overlap and interdependence between the variables, however, it was difficult to evaluate their true relative predictive power. Direct generalization of results to a wide range of target/background conditions would thus involve the assumption of a nearly identical dependence between the variables whenever the technique was applied.

As a result, ridge regression (Hoerl and Kinnard, 1970) was applied to obtain the best estimate of the individual sensitivities and coefficients of the original independent variables. Using this approach, the solution using 15 variables was able to account for 82% of the criterion variance. A reduced set of 7 variables predicted 79% of the criterion variance. The reduced variable set included measured target width, target contrast, target detail contrast, calculated length, ambiguity, heterogeniety, and total brightness elements.

A representative application of the regression equation using the reduced set of seven target/background variables is illustrated in Figure 8, showing a comparison of predicted values with a least-squares curve fit to the actual dynamic acquisition performance data. For the particular target/background case shown, the maximum available range of the target was approximately 36,500 feet. The decrement between this range and the actual acquisition performance curve (.5P = 17,500 ft.) had previously represented an area of largely unpredicted test variability.

| | |
|---|---|
| Factor Analysis | |
| • Correletion Matrix | Pairwise Comparisons With Criterion |
| • Rotated Factor Matrix | Independent Components of Variance – 7 Orthogonal Factors |
| Multivariate Regression | |
| • Linear Stepwise Regression | Quantitative Contribution: |
| 15 Variables | 93% of Criterion Variance |
| Reduced Set (7) | 90% of Criterion Variance |
| • Ridge Regression | Est Independent Sensitivity/Coefficients |
| 15 Variables | 82% of Criterion Variance |
| Reduced Set * | 79% of Criterion Variance (Multiple R = 0.89) |

| | | |
|---|---|---|
| *Reduced Variable Set | Measured Target Width | Ambiguity |
| | Target Contrast | Heterogensity |
| | Target Detail Contrast | Total Brightness Elements (Size Filtered) |
| | Calculated Target Length | |

Figure 8. Target Acquisition Model Prediction of Large-Scale Field Test Data

Continuing target/background classification studies are expected to refine the initial scaling technique, and to verify its application to an expanded range of targets and terrain. Questions also exist as to the extraction of predictive information from different types of premission intelligence/reconnaissance data, as well as the application of prediction techniques to various types of aided-visual and non-visual sensor/display modes. In addition to improved analytical prediction capabilities, however, near-term emphasis will be continued on the identification of a meaningful and economical set of target/background scaling variables that directly impact the decisions of the system engineer and operational field commander.

## CUEING AND PREDESIGNATION AS SEARCH AIDS

Capabilities to quantitatively evaluate and predict visual search performance in a complex and dynamic environment

open the way for the systematic development of performance aid techniques that can best assist the operator with this part of the acquisition task. In addition, methods for objective scaling and classification of target/background conditions will permit the field commander and aircrew to select those operating forces, flight conditions, tactics, sensor modes, and weaponry most likely to result in a successful mission outcome.

As previously shown in Figure 5, targets are rarely acquired at a range and time approximating their first availability to the operator. Rather, target/background conditions more typically result in search requirements that impose severe range penalties. Not infrequently these constraints mean that the target is never acquired at all, even though it has been visually available for a considerable time and range. Such range/time penalties translate themselves into lowered accuracies during rapid target designation, reduced probabilities of successful conversion and weapon delivery, and other decrements that lessen the chances of a successful first-pass target encounter. A critical need thus exists to develop performance aids that significantly assist the operator with the visual search problems imposed by target/background conditions.

Performance aiding concepts vary over a wide range, both as to the basic approach and the manner in which they are employed and mechanized. In general, most involve providing the operator with new items of collateral information designed to speed the search task, or presenting existing information in a more timely and readily usable form. Most have the objective of reducing the amount of visual sampling and processing that must be performed by the operator in real-time, and earlier tests have shown that the most successful techniques for accomplishing this will often be heavily dependent on the particular mission situation, flight conditions, and environment, as well as the availability of collateral data from either on-board or cooperative air/ground sources.

One of the most elementary forms of aiding may be considered to be the development of premission reconnaissance/intelligence data that are closely tailored to the actual conditions under which the target will be encountered. Recent tests indicated that, for certain target/background and mission conditions, such data can result in a 2 to 1 improvement in acquisition performance (Jahns and Gilmour, 1968).

Further improvements have been indicated when such data are presented on a timely inflight basis. Examples, showing a progressively more useful format, include digital display of target/checkpoint coordinates, direct readout of azimuth and time-to-go information, and sequential pictorial display of checkpoints and target lead-in cues that can be directly correlated with the operator's real time visual scene or display.

One extension of these concepts has recently been tested in detail, and serves to illustrate the improvement potentials available (Leininger and Logan, 1971). In this concept, predicted target/checkpoint locations were predesignated to the aircrew via a target window in a head-up display based on onboard navigation system information. Several levels of navigational precision, and predesignation accuracy, were evaluated. Testing consisted of 600 encounters, involving 20 tactical targets, using wide-screen 70mm color motion picture simulation techniques.

Overall results are summarized in Figures 9-11. Figure 9 indicates the performance of all predesignation encounters relative to previously-established unaided performance levels. The upper limits of performance are also indicated by the target availability curve (theoretical

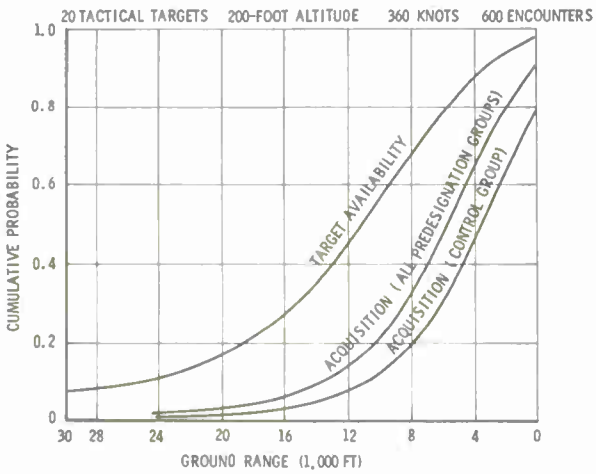



Figure 9. Head-Up Target Predesignation

maximum performance by all aircrews on all target encounters). Performance improvements in acquiring individual targets using predesignation frequently permitted successful target conversion and weapon delivery that would not otherwise have been possible.

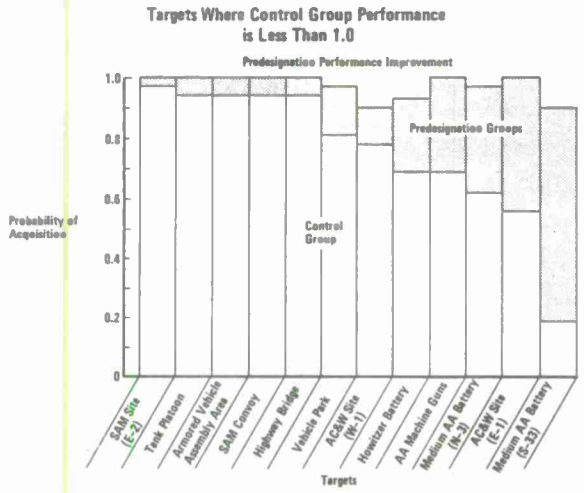



Figure 10. Probability of Acquisition

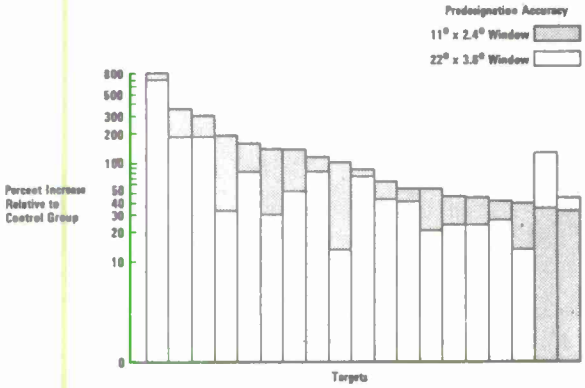



Figure 11. Acquisition Range Improvement Obtained with Predesignation

Figures 10 and 11 show some of the individual target/background effects, and the extent to which the predesignation technique was successful in reducing the operator's search task. Predesignation was observed to be highly effective in reducing missed targets, as reflected in overall acquisition probabilities. Acquisition range improvements were also significant, and quite target specific, with certain target/background situations benefiting considerably more from the use of the predesignation technique.

Further extensions of this aiding concept are currently being developed and evaluated, including the coordinated use of inertial and radar-based predesignation in combination with range-extending concepts such as TISEO. Consideration is also being given to acquisition aiding techniques which provide a smooth interface and transition into fire control and delivery methods using both conventional and standoff weapons.

A wide range of aiding concepts for both day-visual and night/all-weather applications remain to be evaluated. These include the selective filtering of meaningful information from different sensor sources, each of which may be capable of collecting much more data than the operator can process, and the systematic use of cueing data to enhance performance with the most appropriate primary acquisition sensor. Also, sources of cueing information are not limited to onboard systems, but may include cooperative air/ground elements as well.

In a study currently in progress, cueing requirements for a target acquisition system utilizing a primary FLIR display are being investigated. Of particular concern are the resolution and discrimination capabilities of the cueing sensor, and the extent to which inaccurate false alarm cueing data (from whatever source) will be apt to counteract the benefits of cueing on the primary acquisition display.

No single approach to aiding the operator in the target search task is likely to provide an overall solution to the target acquisition problem. Rather, as in the case of weapon-target matching, aiding techniques will have to be developed and applied with a view to the different mission, environment, and system design conditions that apply, and to the particular target/background search situations that result when these conditions are imposed. Analysis, modeling, laboratory research, simulation testing, and field test verification each have an important role in this process.

## REFERENCES

1. Gilmour, J. D., and Emerson, P. L., "Low Altitude, High-Speed Visual Acquisition of Tactical and Strategic Ground Targets,

Part V: A Scheme for Composite Modeling of Parametric Variables and Operational Performance Criteria," D6-2385-5, The Boeing Company, Seattle, 1965.

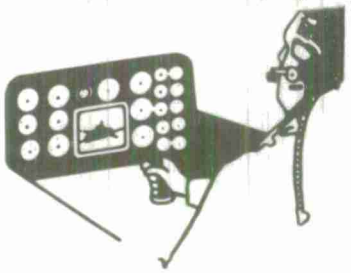
2. Hoerl, A. E., and Kinnard, R. W., "Ridge Regression: Biased Estimation for Nonorthogonal Problems", Technometrics, Vol. 12, No. 1, Feb. 1970, pp. 55-67.

3. Jahns, D. W. and Gilmour, J.D., "Report on Low Altitude Test 4.1, Visual Target Acquisition: Simulation Studies, Extension A," JTF-2-4.1-7-3, Joint Chiefs of Staff, Joint Task Force 2, November 1968.

4. Leininger, W. E. and Logan, A. L., "Airborne Visual Acquisition of Ground Targets - Performance Improvement with Head-Up Predesignation," D162-10310-1, The Boeing Company, Seattle 1971.

5. Wyman, M. J., Gilmour, J. D., and Snyder, H. L., "Report on Low Altitude Test 4.1 - Simulation Studies, Basic Validation", JTF-2-4.1-7-2, Joint Chiefs of Staff, Joint Task Force 2, November 1968 (Confidential).

6. Zaitzeff, L. P., "Target Background Scaling and Its Impact on the Prediction of Aircrew Target Acquisition Performance", Office of Naval Research TR D180-14156-1, The Boeing Company, Seattle, December 1971.



# A DESCRIPTION OF THE TARGET ACQUISITION WORKING GROUP

by

Ronald A. Erickson
Naval Weapons Center
China Lake, California

OFFICE OF NAVAL RESEARCH
TARGET ACQUISITION SYMPOSIUM
NAVAL TRAINING CENTER, ORLANDO, FLORIDA / 14,15,16 NOVEMBER 1972

# A DESCRIPTION OF THE TARGET ACQUISITION WORKING GROUP

By
Ronald A. Erickson
Naval Weapons Center
China Lake, California

## ABSTRACT

This paper describes a newly formed, tri-service Target Acquisition Working Group (TAWG) which has been formed under The Joint Munitions Effectiveness Manuals/Air-to-Surface division of The Joint Technical Coordinating Group for Munitions Effectiveness. The TAWG's scope, organization, and current program are described.

## ORGANIZATION

This paper is a description of a recently-formed, tri-service Target Acquisition Working Group (TAWG). The purpose of the paper is to make the target acquisition community aware of The TAWG's existence, and to improve communication and coordination of work in target acquisition. Being a tri-service group, we solicit aid from all thirds. The TAWG is organized under The Joint Munitions Effectiveness manuals for Air-to-Surface (JMEM/AS). Perhaps some of you are familiar with the munitions effectiveness manuals that are produced by The JMEM. The Chairman of JMEM/AS is Mr. Thomas Christie from Eglin Air Force Base, and the Co-chairman is Dr. Margarite Rogers from NWC, China Lake.

The JMEM/AS is a division of The Joint Technical Coordinating Group for Munitions Effectiveness (JTCG/ME) which is chaired by Dr. Joseph Sperrazza. Dr. Sperrazza, who is also the technical director of The Army Material Systems Analysis Agency, reports on JTCG activities directly to The Joint commanders. The hierarchy is shown partially in Figure 1.

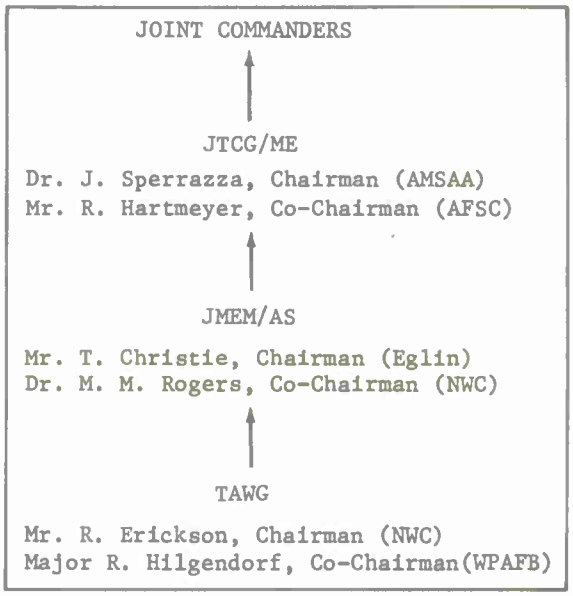



FIGURE 1. TAWG Line-of-Responsibility

## OBJECTIVE

The JTCG/ME is responsible for producing manuals which provide all DOD elements a uniform basis for:

- Weapons employment planning and use
- Determination of munitions requirements
- Evaluation of new weapon concepts
- Stimulation of R & D efforts.

The TAWG objectives within this framework are to quantify target acquisition performance, and to develop a means of predicting performance. The only current bounds on The TAWG scope are limitations to

(1) Air-to-ground target acquisition (TA)
(2) real-time operations.

We are assuming that there will be at least two types of customers for our output: The operating forces, who would like to know the level of performance to be expected for any given mission, and The R & D community who are interested in estimating the effectiveness of existing, prototype, and hypothesized systems.

## SCOPE

Real-time, air-to-ground target acquisition (TA) can involve direct vision, sensor/display systems, and optical systems. These, in turn, can be broken down into subdivisions; e.g. sensor/displays include television, infra-red, and radar systems. As far as I have been able to determine, The TAWG will be dealing with all of these, although funding and manpower levels are limiting us primarily to TA via direct vision at the moment.

Since The TAWG is a tri-service group, we will be dealing with a variety of aircraft: OH-58 and AH-56 helicopters, the A-8 Harrier, The OV-10, A-7's to F-4's, F-14's, and the new AX. The visibility from these aircraft, mission types, and speed and altitude regimes must be included among our target acquisition parameters.

Any mission involving target acquisition can be broken into segments; those segments with which The TAWG is primarily concerned are (1) target area recognition, (2) search for the target, and (3) detection/recognition/identification of the target. Although factors like reconnaissance, pilot training, pilot briefing, target designation, and conversion to attack are in mission success, The TAWG is hoping to bite off only the 3 items listed above.

## METHODOLOGY

Several methodologies have been used in the past to produce estimates of target acquisition performance: mathematical modeling, simulation (e.g. with photos or terrain models), field testing, and use of combat data. The TAWG intends to use all appropriate methods and has, in fact, partially organized around methodology type.

## TARGET ACQUISITION DEFICIENCIES

As far as I can tell, the reasons for The TAWG's existence are (1) the paucity of usable data that can be applied to mission planning and R & D efforts, and (2) the need for a tri-service group to answer TA questions which arise. Some of the deficiencies which we have identified are listed in Table 1 below. These deficiencies were used to define some of the candidate TAWG projects shown in Table 2. Only some of these are being pursued at this time.

TABLE 1. SOME TARGET ACQUISITION DEFICIENCIES

MATH MODELING INADEQUATE
MOST MATH MODELING NOT VALIDATED
CONTEXT (IMBEDDEDNESS, CONSPICUOUSNESS) NOT QUANTIFIED
PERTINENCE OF LAB STUDIES TO FIELD PERFORMANCE NOT KNOWN
SEARCH METHODS NOT KNOWN
INDIVIDUAL DIFFERENCES CANNOT BE ACCOUNTED FOR OR EXPLAINED
MOST FIELD TESTS POORLY REPORTED
MOST FIELD TESTS POORLY STRUCTURED
FIELD TESTS NONCOMPLEMENTARY (DIFFICULT TO SUMMARIZE AND COMBINE)
IMAGE QUALITY DESCRIPTORS NOT STANDARDIZED
NO STANDARDIZED METHOD OF SYSTEM ADJUSTMENT AND CALIBRATION
TARGET/BACKGROUND SIGNATURES NOT MEANINGFULLY QUANTIFIED OR DEFINED

TABLE 2. CANDIDATE TAWG PROJECTS

TERRAIN/GROUND TARGET CLASSIFICATION/QUANTIFICATION
- ·IMBEDDEDNESS
- ·CONSPICUOUSNESS
- ·CAMOUFLAGE

DISPLAY QUALITY/PERFORMANCE
- ·STANDARDIZE MEASURES OF QUALITY
- ·RELATE QUALITY MEASURES TO OPERATIONAL PERFORMANCE

MATH MODEL REVIEW AND SYNTHESIS
SHIP ACQUISITION PERFORMANCE
- ·USE OF TACTICS
- ·USE OF SENSORS
- ·RECOGNITION CRITERIA & PERFORMANCE

GEOGRAPHIC ORIENTATION IN TARGET AREA
- ·WHERE, WHEN TO SEARCH FOR TARGET
- ·EXPECTATION OF TARGET PRESENCE

USE OF FAC's
- FAST FAC's; CAN WE USE FAC's IN SITUATIONS OTHER THAN SEA?

USE OF FLARES
- NUMBER, LUMINANCE, PLACEMENT, ALTITUDE
- AIRCRAFT TACTICS

INITIAL TAWG TASKS

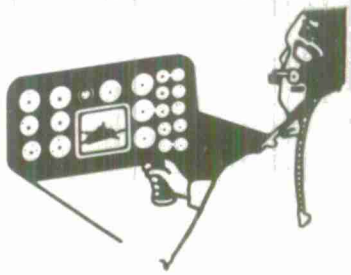
Some of the major current TAWG projects are shown below with some of the work components indicated.

1. MATH MODEL REVIEW AND SYNTHESIS
   - DEVELOP EVALUATION CRITERIA
   - REVIEW AND CLASSIFY MODELS
   - EVALUATE FOR INPUT AVAILABILITY, VALIDITY, ASSUMPTION CONTENT, OUTPUT FORMAT, COMPUTABILITY
   - RECOMMEND MODEL(S) FOR TAWG USE
   - RECOMMEND MODEL DEVELOPMENT/REFINEMENT (e.g. SENSITIVITY ANALYSES)
   - RECOMMEND PROGRAM FOR PROVIDING OR OBTAINING REQUIRED MODEL INPUTS
2. USE OF FLARES
   - CONTINUE FLARE RESEARCH
   - REVIEW ALL FLARE PROGRAMS
   - SUMMARIZE FINDINGS TO DATE
   - PREPARE FLARE USE MANUAL
3. FORWARD AIR CONTROLLER
   - PROBLEM DEFINITIONS, TASK ANALYSIS, IDENTIFY CRITICAL TASKS
   - PRELIMINARY ESTIMATION OF TA PERFORMANCE
4. TERRAIN/TARGET CLASSIFICATION
   - TERRAIN TAXONOMY (PHOTO STUDY)
   - TERRAIN/TARGET COMBINATIONS
     IMBEDDEDNESS
     CONSPICUOUSNESS
   - INDICES OF TARGET ACQUISITION DIFFICULTY
5. FIELD DATA SUMMARY
   - DEVELOP EVALUATION CRITERIA
   - EVALUATE REPORTS OF TA FIELD TESTS
6. SUMMARIZE TERRAIN MASKING DATA
   - LITERATURE SEARCH
   - DEVELOP EVALUATION CRITERIA
   - EVALUATE REPORTS
   - SUMMARIZE/SYNTHESIZE DATA
   - IDENTIFY DEFICIENCIES

The work is being carried on under the cognizance of the subgroup chairman shown in Table 3. We are currently working under the assumption that the TAWG will exist for at least 2 or 3 years, chipping away at assignments as funding and staffing permit. One of the larger areas which we have not yet addressed is that of TA aids. This whole area of sensors, displays, cuing devices, etc. is ahead of us.

TABLE 3. TAWG SUBGROUP CHAIRMEN

- TAWG CHAIRMAN —— RONALD ERICKSON, NWC
- MATHEMATICAL MODELING —— DR. H. H. BAILEY, RAND
- ANALYTIC STUDIES —— DARRYL THORNTON, EGLIN AFB
- LABORATORY SIMULATION —— RONALD BRUNS, NMC
- FIELD TESTING —— PAUL AMUNDSON, NWC
- FORWARD AIR CONTROLLER —— LCOL ED WAGGONER, BROOKS AFB
- FLARES —— MAJOR BOB HILGENDORF, WPAFB



# THE EFFECT OF IMAGE QUALITY ON TARGET RECOGNITION

by

Leon G. Williams, PhD, and
Carl P. Graf
Honeywell Research Labs
St. Paul, Minnesota 55113

OFFICE OF NAVAL RESEARCH
TARGET ACQUISITION SYMPOSIUM
NAVAL TRAINING CENTER, ORLANDO, FLORIDA / 14,15,16 NOVEMBER 1972

# THE EFFECTS OF IMAGE QUALITY ON INFRARED TARGET RECOGNITION*

Leon G. Williams and Carl P. Graf
Honeywell Research Labs
St. Paul, Minnesota 55113

The study objective was to determine the relative importance of four image quality variables, lines/target, modulation transfer, signal-to-noise ratio, and target size on target recognition in a static display. The study's conclusions are supported by four factors: First, a realistic target recognition task was used -- 16 examples of each of 10 infrared target categories were obtained for subsequent processing. Second, a high degree of control over the stimuli was obtained by digital image processing techniques. Third, a large volume of experimental data were obtained -- 18 subjects produced data from 15,108 recognition trials. Fourth, linear regression techniques were applied to the data to determine the relative importance of the several image quality variables.

## Method

The initial infrared pictures of military targets (tank, truck, jeep, gun, etc.) were produced at the Honeywell Proving Grounds in Minnesota. An IR sensor in the 1.0-5.5 micron range displayed the thermal images on a CRT for a range of temperature conditions (night, day, target hot, target cold). Photographs were made of these images. The desired levels of image degradation were achieved with the digital image processing facility.

Each target transparency (1 x 2 inches) was written on magnetic tape as 250 records, each record containing 500 bytes, by a drum scanning technique. The main image degradation was achieved by digital computer processing. First, a Gaussian point spread function was applied to create the specified level of modulation transfer. Then, Gaussian noise was added to each sample picture element. Finally, the line scan structure was created. The output tape was made into a 0.5 x 1.0 inch transparency by a drum writer. A total of 1120 transparencies were made.

The experimental set-up consisted of a modified 35mm slide projector, a quality rear projection screen, a timer and printer, and a logic controller for the apparatus. In a trial a transparancy was projected at one of four magnifications to achieve mean target heights ranging from 10 to 80 arc minutes of visual angle. The subject had five seconds to make a recognition response. His response and time were recorded.

## Design

The study consisted of two parts. In the first part, initial evaluation, the objective was to determine the degradation levels to be used for the main experiment. This was accomplished by degrading a set of four pictures at all levels of interest and inspecting the output transparencies. The main outcome of this effort was the decision to use more noise and poorer MTF for those pictures having more scan lines.

In the main experiment there were 224 degradation conditions defined by four levels (with an exception) of each of the

*This work has been sponsored by the Air Force Avionics Laboratory at Wright-Patterson AFB under Contract No. F33615-72-C-1303.

four main image quality variables. The nominal ranges of each variable were as follows: 1) lines/target (3 to 24); 2) MTF (0.08 to 0.47 modulation at one-half the cutoff frequency); 3) signal-to-noise (1:1 to 12:1); 4) target size (10 to 80 arc minutes). There were 18 subjects, 6 with considerable experience with FLIR imagery. All subjects saw the total set of 1120 transparencies, 280 during a three hour training session and the remaining 840 during the experimental trials.

Precise measurements of size and density were made on each of the original target transparencies. There were two reasons for doing this: To define the values of the degradation and to determine which target characteristics other than the four main image quality variables were related to recognition. The following measurements were made on each target: Height, width, minimum dimension, maximum dimension, area on the transparency, perimeter, mean edge contrast, mean density, background density maximum picture density, and minimum picture density. The target measures were combined with the measures of image quality to provide a variety of combined measures. Two examples are target width in terms of number of scan lines and signal-to-noise ratio.

A measure of resolving power was made for each degradation level using the Air Force Resolution Chart. Resolving power was defined by the minimum size 3-bar pattern which could be visually separated at that degradation condition. It was found that resolving power depended almost entirely on the number of lines/target and target size (i.e., magnification).

Results

Data from 15,108 trials were analyzed. The overall recognition accuracy was 58 percent; the mean response time was 2.75 seconds. The correlation of response time and accuracy was -0.93. Although the experienced subjects recognized about as many targets as the students, they were about one second slower. Accuracy was much greater for some targets, such as the tank or truck, than others.

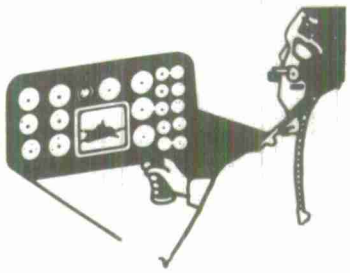
The main result of the regression analysis was that lines/target and noise in linear combination accounted for over 60 percent of the variance in both time and accuracy for the 224 degradation conditions. Resolving power, however, was a better predictor of performances than lines/target. Resolving power and noise together accounted for 78 percent of the variance for the 224 degradation conditions.

What target properties make it easy or hard to recognize? Two kinds of variables were studied, target size, and target-background contrast. If recognition was predicted from target size alone, then size was best measured in terms of resolving power. Specifically, target perimeter, expressed in terms of resolving power, was the best single predictor of performance. This was similar to John Johnson's widely accepted measure, namely, target height measured in terms of resolving power. Target size, expressed in terms of visual angle was a surprisingly poor predictor. Target-background contrast was measured better by mean edge contrast than by the simpler overall contrast measure. In summary, recognition was best predicted by the following variables: resolving power noise, target perimeter, and edge contrast.

Recommendation

The study results strongly suggested that future effort be concerned with a more precise understanding of resolving power. Specifically, how do resolving power and recognition depend on the primary aspects of the sensor-display line scan structure; i.e., sensor scanning aperture--size, shape, and spacing; and display writing aperture--size, shape and spacing.

THIS IS AN ABSTRACT OF THE PAPER PRESENTED AT THE SYMPOSIUM.
FOR FURTHER INFORMATION, THE AUTHORS MAY BE CONTACTED.



# AN ANALYSIS OF RADAR TARGET ACQUISITION PERFORMANCE EFFECTS ON SURVEILLANCE AND DIRECTION EFFECTIVENESS THROUGH SIMULATION

by

Donald A. Topmiller, PhD
and
Wayne L. Martin

Aerospace Medical Research Laboratory
Aerospace Medical Division
Air Force Systems Command
Wright-Patterson Air Force Base, Ohio

OFFICE OF NAVAL RESEARCH
TARGET ACQUISITION SYMPOSIUM
NAVAL TRAINING CENTER, ORLANDO, FLORIDA / 14,15,16 NOVEMBER 1972

# AN ANALYSIS OF RADAR TARGET ACQUISITION PERFORMANCE EFFECTS ON SURVEILLANCE AND DIRECTION EFFECTIVENESS THROUGH SIMULATION

Donald A. Topmiller, Ph.D.
and
Wayne L. Martin

Aerospace Medical Research Laboratory
Aerospace Medical Division
Air Force Systems Command
Wright-Patterson Air Force Base, Ohio

## ABSTRACT

This paper summarizes studies done in command/control simulations of surveillance functions where the operator is required to detect threats through an automated digital radar system and of weapons direction functions. The effectiveness of the operator's performance depends largely on the ability of the operator to acquire a nonimaging and digitally displayed target and to identify what he perceives as being different from nontarget returns using the time course history of the returns. He must then assign this information to a computer programmed to automatically track the progress of the target across the surveilled area. The operator may perform this coordinate position identification and subsequent track initiation with a light pen or through cursor positioning with a track ball, joystick, or other means. The weapons direction function must be performed as effectively as the surveillance function in order to achieve maximum defense capability.

## INTRODUCTION

Recognizing that the primary purpose of this symposium is to appraise the state-of-the-art in acquiring targets which have "image" characteristics, this paper is intended to present a wholly different problem in target acquisition, viz., that of acquiring targets through nonimaging sensors (i.e., digitalized radar). The problem of identifying target characteristics, as well as problems associated with counteracting techniques designed to purposely distort these characteristics (e.g., camouflage for imaging targets and ECM for electronically sensed targets), is more universal than that associated with digital radar systems. However, the main intent of this paper is to explore target acquisition and directing capabilities in "automated" digital radar systems.

To discuss this area and its associated experimental findings from a two-year program of computer-based simulation first requires some conceptual distinctions.

Behaviorally, including both visual mechanisms and the more general realm of perception, imaging capability has a greater experiential and experimental data base to draw upon than nonimaging information structures and displays. After all, the giant strides in electronic and computer technologies have only occurred within the last 30-40 years to make this possible. Biologically, man has evolved into an astute visual perceptual sensor of his environment; viz., all information transformation is biologically and psychologically determined. Alternately, technology growth has enabled man to be a secondary detector and perceiver of his environment, wherein the information

has been physically sensed [RF (radio frequency) energy, IR (infrared) energy, etc.], and, of necessity, filtered, transformed and displayed. Albeit these filters and transformation schemes were developed by man himself, frequently they are not designed to maximally exploit the operator's information receiving-processing capabilities. With analog radar and infrared sensors, there are built-in perceptual cues that are not necessarily remote from the "true" image. However, with digital radar, the sheer process of digitalization of the sensed data, normally driven by temporal considerations such as antenna scan rate, provides a greater challenge to the engineering psychologist responsible for developing the research necessary to define optimal information, transformation, structure and display techniques.

Turning now, from the information sensing and transformation stages of the target acquisition process, to the command-control implications, higher order psychological functions are demanded from the operator: once the target has been acquired, what resources should be committed to most effectively eliminate the threat. A very limited facet of this problem has been addressed through recent Airborne Warning and Control (AWACS) weapons direction simulations in our laboratory. A more extensive aspect of the "inferential process" involved in tactical decision-making was treated in a six-year program at Ohio State University Research Foundation and has been extensively documented elsewhere (Howell, 1967; Howell and Gettys, 1968).

## SOME CONCEPTS IN DIGITALIZED SURVEILLANCE SYSTEMS

### Radar History

One of the often professed advantages of the "automated" surveillance radar system is the ability to sense, store and play back radar returns from several sensing antennae. The "storing" function allows the systems to be designed such that the actual history of the scans can be played back at any rate of which the computer is capable. Recent research (Scanlon et al, 1971) has shown the advantage of time-compressed radar returns for display purposes since the targeting returns can be more readily distinguished from nontarget information, i.e., clutter.

### Tracks-Radar Trails

The actual "playback function" in a digital radar system creates a series of coherent blips on the scope which can be defined as a radar trail. Returns that are "canned" to represent false reports may have any degree of temporal/spatial similarity to data that have been "canned" to represent the timed history of actual aircraft. The "canning" can also define the start and stop positions of the trail, velocity of the trail, type of return (i.e., Mark X, search or false reports) and time of entry for each trail.

### Clutter

Since any RF energy reflecting surface (such as automobiles, bridges, TV towers, etc.) can produce a return in a digital system, without exotic filtering techniques the system cannot discriminate noisy returns or clutter from true target returns. On more traditional analog radar systems, the return patterns can give perceptual cues as to distinguishing characteristics of clutter versus targets which enable the surveillance operator to exploit this information. In the simulations of digital radar systems, the clutter is typically generated by some statistical distribution; in several of our studies, we mixed uniform and exponential distributions. This method provided a realistic distribution of clutter that was unevenly distributed over the surveillance area and contained "clumping".

### Blip/Scan Ratios

Because of other noisy sources, including atmospheric degradation, intentional countermeasures, etc., not every energy pulse from a scan of the antenna will have an associated return. The ratio of the emitted scans to the sensed returns is referred to as the blip/scan ratio. This parameter can be varied in our simulations (typically from .5 to .9) and represents a true capability to systematically degrade the quality of the information available to the surveillance operator.

Track Failure

The tracking algorithms, in general, are subject to failure due to many influences such as maneuvering aircraft, crossing aircraft, blip/scan, clutter and radar position errors. No matter how effective the algorithm is, certain radar data systems will cause a tracking failure. The failure occurs when the predicted position is caused, by one or more of the above influences, to be at the wrong coordinates with respect to the radar data from the aircraft being tracked. When this happens, the tolerance can be opened up.

## THE OPERATOR AS A THREAT DETECTOR

Single Operator Simulations

Early concerns in advanced, automated digital radar systems dealt with the number of targets that a single operator could be expected to detect and assign to the computer for automatic tracking. The two primary threat/radar environment parameters which were thought to contribute most significantly to the operator's ability to perform the track initiation and maintenance functions were clutter and the rate at which the penetrators entered the display area.

Robert G. Mills, from our group, was principally responsible for developing the single-operator simulation studies which are more thoroughly discussed in two AMRL technical reports (AMRL-TR-70-103 and AMRL-TR-71-76).

The simulation facility consists of an IBM 360-40H computer system which drives four IBM 2250 cathode ray tube (CRT) graphics terminals which are used as the primary control/display units. The terminal has a CRT display surface of 144 square inches (12 in. x 12 in.). The CRT is coated with P-7 phosphor which persists on the average of 400 ms after excitation. The terminal's light pen and programmed function keyboard (PFK) consist of 32 response keys for operator communication with the computer.

Diagram A is a representation of a time exposure photograph taken over two radar updates (six scans are shown) during an early 20-second period of a mission. A number of targets are shown, several of which have been initiated.

The subject's task was to monitor his surveillance area and to perform track initiation and maintenance functions. The initiation function required the subject to light pen any one of five historical returns suspected of representing a target. Following this action, the subject activated a response key labeled NT (new track). The maintenance function was performed in the same manner, except that the subject pressed a response key labeled OT (old track) instead of NT when he detected the disassociation of a track with a signature.

As mentioned previously, two independent variables (clutter and target introduction rate) were manipulated in the simulation. The levels of clutter were .016 pieces/square nautical mile (C1, 90 total pieces/scan), .032 pieces/square nautical mile (C2, 180 total pieces/scan), and .064 pieces/square nautical mile (C3, 360 total pieces/scan). The levels of target introduction rate were 1, 2, 3, and 4 targets/minute. The mission took approximately 50 minutes of real time, and subjects completed three missions/week over a two-month period.

One of the findings of this simulation is the positive relationship between the two surveillance functions of initiation and maintenance. This relationship is not necessarily surprising (see Figure 1) and perhaps could have been predicted from basic research on attention (vigilance) behavior. In Figure 1 the parabolic function demonstrates that the more effort the subject devoted to maintaining targets and keeping his display cleaned up, the higher the probability of initiating new targets as they entered the surveillance area.

Figure 2 shows the effects of the two independent variables of clutter and target introduction rate. The surprising result is that clutter appears to have very little effect on probability of correct initiation and was not a significant main effect in an analysis of variance, whereas target introduction rate was highly significant and accounted for

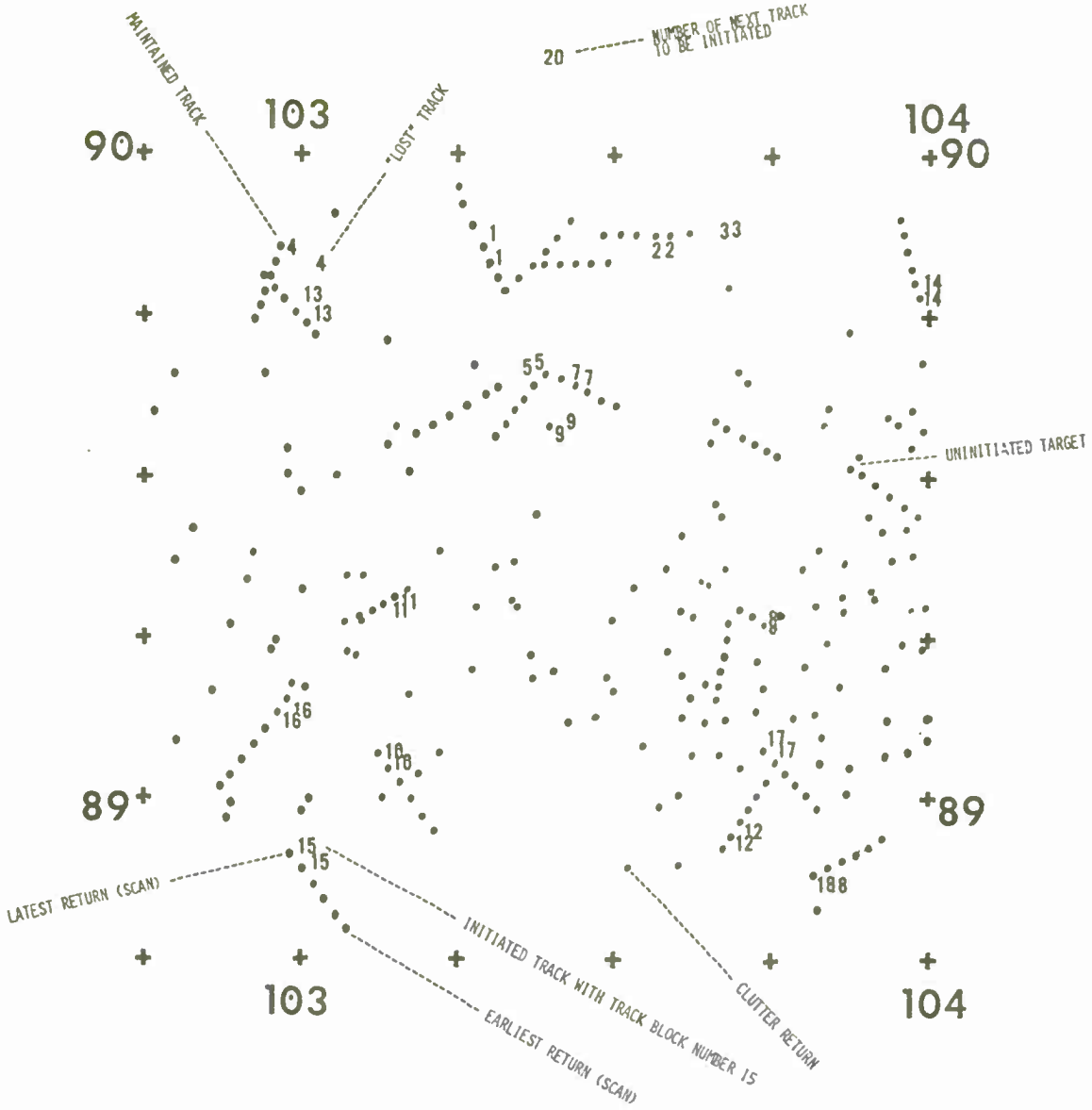



Diagram A. Representation of a Time Exposure Photograph of the CRT Display Showing One Display Update During a 20-Second Mission Simulation Period

the majority of variance. Evidently the visual-perceptual discriminations of the operator are better than commonly thought by the surveillance community.

Mills summarizes his conclusions and recommendations as follows:

- Operator performance is degraded particularly under target introduction rates of 3 and 4 per minute. Until other alternatives are investigated two operators should perform the surveillance tasks when introduction rates reach these levels.

- Operator performance is relatively unaffected by clutter density under the time compression display format, however, further investigations should be performed to determine whether or not this conclusion holds under less idealized radar display conditions.

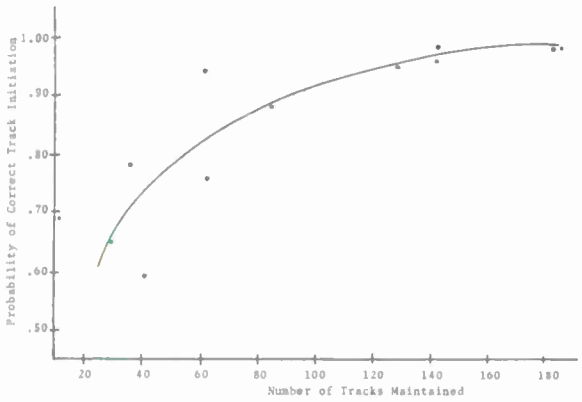



Figure 1. Probability of Correct Track Initiation as a Function of Number of Tracks Maintained for Two Mission Conditions (Introduction Rate = 4, Clutter Density = 0.16 and 0.32). Mission Time Covered = 50 Minutes Number of Subjects = 6.

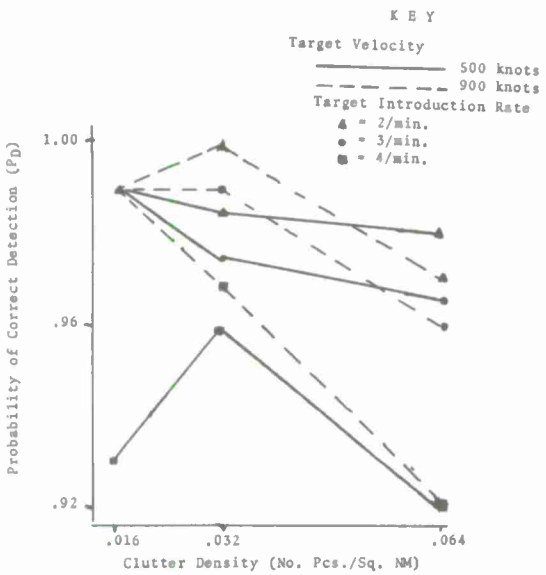



Figure 2. Summary Data for First Two Single-Operator Simulations of Command/Control Surveillance Functions. Data are Averages Across Operators.

Multi-Operator Simulations

This program provided a surveillance system simulation in which up to four operators could perform the tasks of track initiation and maintenance under various digital display parameters.

Some of the parameters were fixed but could be made to assume any value appropriate to the research problem at hand. These fixed parameters were established at "can" time and could be manipulated in any fashion to provide the following variable characteristics: start and stop position of trail, velocity of trail, type of return (i.e., Mark X, search, or false reports), and time of entry for each trail.

In combination, these parameters could be manipulated as independent variables in an experiment to provide variable introduction rate, variable enemy attack strategies, any mix of trail velocities, a broad range of trail lengths, false report density, any blip/scan ratio and radar return errors could be manipulated mathematically/statistically.

Returns that are "canned" to represent false reports may have any degree of temporal/spatial similarity to data that have been "canned" to represent the timed history of actual aircraft.

Once the "canned" portion of the program has been constructed, a large portion of what the operator saw (and thus a good deal of his work load) was determined.

The values of other relevant parameters could be designated at the time a particular experimental condition is run. Given the identical "canned" program (or set of programs), the additional independent variables which can be manipulated are as follows:

- probability of track failure - Mark X data
- probability of track failure - Search data
- heading tolerance permitted on initiate & reinitiate actions - Mark X
- heading tolerance permitted on initiate & reinitiate actions - Search
- drift threshold - used to hold off drift-in logic until operator has a chance to act
- trouble track indicator - number of scans a track may drift before it is considered a trouble track
- correlation specification for track-to-aircraft correlation - this

defines the size of the area around the tail of a drifting track within which search returns will become correlated with the track

- entire geographic background display including boundary lines, strategic location designators, and latitude/longitude markers

Several display options were provided through program functions keys. The operator (subject) may omit or include any of the following categories of data:

- uncorrelated history data for search radar
- correlated history data for search radar
- uncorrelated present position data for search radar
- correlated present position data for search radar
- uncorrelated history data for Mark X radar
- correlated history data for Mark X radar
- uncorrelated present position data for Mark X radar
- correlated present position data for Mark X radar
- trouble tracks
- good tracks
- boundary lines
- strategic locations
- latitude/longitude markers

Figure 3 shows the layout of PFK's.

The operator also provided a heading prediction and expanded the display via the PFK's as indicated. An initiate action, for example, required depressing the initiate key, light penning the most recent of a series of returns and depressing the appropriate heading prediction key.

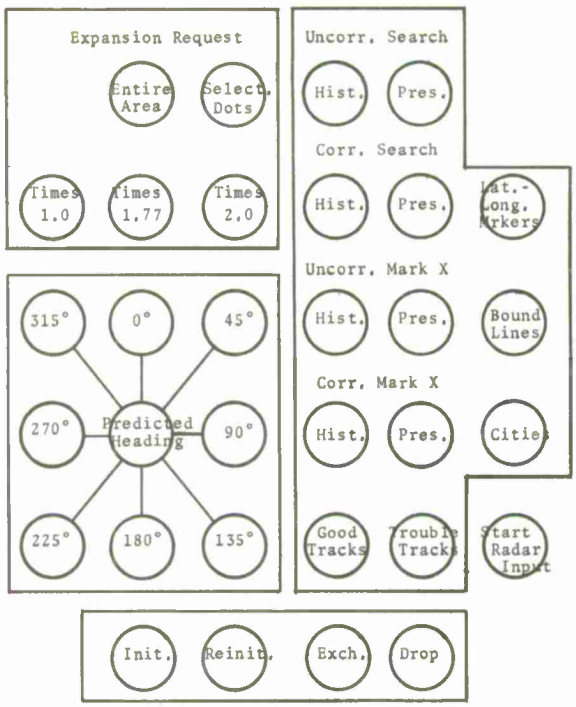



Figure 3. Programmed Function Key Layout

All operators actions, whether they were good or bad (and if bad, the cause), and every event occurring throughout a simulated set of radar data were recorded on a scan-by-scan basis. The status of each track (i.e., drift or follow mode) was also recorded. At the conclusion of an experimental session, the frequency of each event category was compiled. The definitions of these categories are shown in Table I.

Performance measures were also evaluated automatically at the end of each experimental session. These measures, together with their derivation, are provided in Table II. Measures 1-6 relate to operator actions and represent several ways of describing what the operator was doing and how well he was performing.

Measure 7 was taken as the overall surveillance system performance. "Good Tracking" was the case when a track was followed properly. When a track was in drift status, "Bad Tracking" was said to be the case.

TABLE I

DEFINITION OF EVENT CATEGORIES

| | | |
|---|---|---|
| Scans | = | Scans for thie run |
| EN | = | Aircraft entered |
| SC | = | Scale Change |
| IG | = | Initiate Good |
| IB - AI | = | Initiate Bad - Already Initiated |
| RG | = | Reinitiate Good |
| RB - WT | = | Reinitiate Bad - Wrong Track aseociated with trail |
| IB - AL | = | Initiate Bad - Aircraft Landed |
| RB - AL | = | Reinitiate Bad - Aircraft Landed |
| IB - BH | = | Initiate Bad - Bad Heading prediction |
| RB - BH | = | Reinitiate Bad - Bad Heading prediction |
| IB - AF | = | Initiate Bad - Already being Followed |
| RB - AF | = | Reinitiate Bad - Already being Followed |
| RB - TD | = | Reinitiate Bad - Track already Dropped |
| DG | = | Drop Good |
| DB - SA | = | Drop Bad - Track Still Aseociated with a trail |
| DB - NT | = | Drop Bad - No Track preeent |
| EG | = | Exchange Good |
| EB - NO | = | Exchange Bad - Not the Original aircraft |
| EB - NT | = | Exchange Bad - One aircraft has No Tract |
| EB - MT | = | Exchange Bad - One aircraft was Mark X (Ten) |
| DI | = | Drift In |
| DO - ST | = | Drift Out - Statietics |
| DO - AL | = | Drift Out - Aircraft Landed |
| TT | = | Trouble Track Status |
| SL | = | System Limits |
| IB - CL | = | Initiate Bad - Clutter |
| RB - CL | = | Reinitiate Bad - Clutter |
| DG - CL | = | Drop Good - Clutter |
| EB - CL | = | Exchange Bad - Clutter |
| IB - AF Ignored | = | Thoee Initiate Bad - Already being Followed actions that occur within 2 ecane of an IG |
| RB - AF Ignored | = | Those Reinitiate Bad - Already being Followed actione that occur within 2 scans of an RG |

The purpose of this study was to evaluate operator and surveillance system performance at various trail lengths, across blip/scan ratios for 1, 2, and 3 man teams. The effort was performed to settle the question concerning how many history returns (history + present = trail length) must be stored and displayed to the operator.

TABLE II

PERFORMANCE MEASURES

1. Mean Firet Good Initiate Time =

$$\frac{\sum_{IG}^{EN} (\text{SCAN \# IG} - \text{SCAN \# EN})}{\text{Total \# IG}} \times 10.0$$

2. Percent of Initiate Actione That Were Correct =

$$\frac{\text{Total \# IG}}{\text{Total \# IG + IB-AI + IB-AL + IB-BH + IB-AF + IB-CL}} \times 100$$

3. Percent of Aircraft Correctly Initiated =

$$\frac{\text{Total \# IG}}{\text{Total \# EN}} \times 100$$

4. Percent of Reinitiate Actions That Were Correct =

$$\frac{\text{Total \# RG}}{\text{Total \# RG + RB-WT + RB-AL + RB-BH + RB-AF + RB-TD + RB-CL}} \times 100$$

5. Percent of Good Maintenance Actions Taken =

$$\frac{\text{Total \# RG + DG + EG}}{\begin{array}{c}\text{(Total \# RG + DG + EG + RB-WT + RB-AL + RB-BH + RB-AF + RB-TD +} \\ \text{DB-NT + DB-SA + EB-NO + EB-NT + EB-MT + RB-CL + EB-CL)}\end{array}} \times 100$$

6. Gross Merit = $\left[\text{(Good - Bad) / (Good + Bad)}\right] \times 100$

Where: Good = IG + RG + DG + EG + DG-CL

Bad = IB-AI + RB-WT + IB-AL + RB-AL + IB-BH + RB-BH + IB-AF + RB-AF + RB-TD + DB-NT + DB-SA + EB-NO + EB-NT + EB-MT + IB-CL + RB-CL + EB-CL

7. Overall Percent of Scans With Good Tracking =

$$\frac{\text{Total \# Scans With Good Tracking}}{\begin{array}{c}\text{(Total \# Scans With Good Tracking +} \\ \text{Total \# Scans With Bad Tracking)}\end{array}} \times 100$$

Previous studies using a BUIC III type simulation (Pearson, 1972) showed that a trail length of 7 was adequate for tracking through noise. Blip/scan, however, was held at unity in these studies.

The mission parameters were:

- Mission duration of 28 minutes
- Aircraft entered the 200 x 200 nautical mile surveillance area at the average rate of 2/minute ($\sigma$ = .75/minute) with a mean velocity of 500 nautical mph ($\sigma$ = 150 nautical mph)
- Radar return error was distributed as a Rayleigh with $\sigma$ = 1 nautical mile
- Probability of track failure = .13
- False report rate (or clutter) was 2/scan across each of the trail lengths, resulting in 10, 14, and 18 total pieces distributed as a uniform bivariate in X and Y display coordinates

The experimental design was as follows:

- The design employed a 3x3x3 factorial with 4 replications.

- Factor levels:

  - Trail length was manipulated at values of 5, 7 and 9.

  - Blip/scan was varied at levels of .5, .7 and .9.

  - Teams of 1, 2 and 3 operators performed track initiation and maintenance tasks.

  - Added experimental control was provided by using the same trained operators who participated in earlier studies. An original 4-man team was broken down into 1-, 2- and 3-man combinations such that counterbalancing was maintained in so far as possible with respect to an individual operator's contribution to performance by each team size. To further enhance control, all the data required for four replications across group size within any particular trail length X blip/scan combination was collected on a single day. The only restriction was that no operator ran more than two 28-minute sessions in a row and each had at least a 28-minute break between all other sessions.

Analyses of variance were performed on each of the seven performance measures. Summaries of these analyses are provided for only the Mean First Good Initiate Times and Overall Percent Good Tracking and are shown in Tables III and IV. Likewise, only the figures plotting means for First Good Initiate Time and Overall Percent of Scans with Good Tracking are presented. Figures 4 through 6 can be overlayed and the lack of an effect of trail length on Mean First Good Initiate Time easily observed. Likewise, Figures 7 through 9 portray essentially the same nonsignificant effect of trail length. Blip/scan had a statistically significant effect (at least at the .05 level) for every performance measure, while trail length was not significant for any of the measures. Group size had a significant effect on all measures except Percent of Initiate Actions That Were Correct. The only interaction found to be significant was on the Overall Percent of Scans With Good Tracking measure for trail length versus blip/scan. Increasing the trail length does enhance overall surveillance system performance at lower blip/scan levels.

TABLE III

ANALYSIS OF VARIANCE FOR MEAN FIRST GOOD INITIATE TIME

| Source of Variance | df | MS | F |
|---|---|---|---|
| A (Group Size) | 2 | 2595.14 | 24.751* |
| B (Trail Length) | 2 | 60.74 | .579 |
| C (Blip/Scan) | 2 | 8960.48 | 85.459* |
| A x B | 4 | 35.25 | .336 |
| B x C | 4 | 208.59 | 1.989 |
| A x C | 4 | 56.61 | .539 |
| A x B x C | 8 | 20.51 | .196 |
| Error: Within Treatments | 81 | 104.85 | |
| Total | 107 | | |

* $p < .001$

TABLE IV

ANALYSIS OF VARIANCE FOR OVERALL PERCENT OF SCANS WITH GOOD TRACKING

| Source of Variance | df | MS | F |
|---|---|---|---|
| A (Group Size) | 2 | 201.64 | 19.714* |
| B (Trail Length) | 2 | 23.92 | 2.339 |
| C (Blip/Scan) | 2 | 504.71 | 49.347* |
| A x B | 4 | 5.28 | .516 |
| B x C | 4 | 40.00 | 3.912* |
| A x C | 4 | 18.75 | 1.833 |
| A x B x C | 8 | 4.70 | .459 |
| Error: Within Treatments | 81 | 10.23 | |
| Total | 107 | | |

* $p < .01$

The seemingly discrepant observation for blip/scan = .5, trail length = 7 in Figure 8 is probably due to the fact that this particular combination was the first cell in which performance data were collected in this experiment. Subjects had undergone considerable training at the identical mission parameters but apparently still had not reached asymptotic performance when this study was initiated. Analysis of subsequent data indicated that little, if any, improvement in performance occurred as a function only of practice.

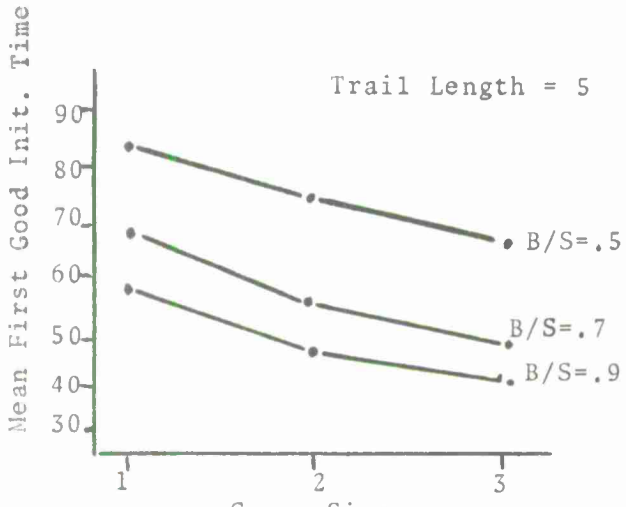



Figure 4.

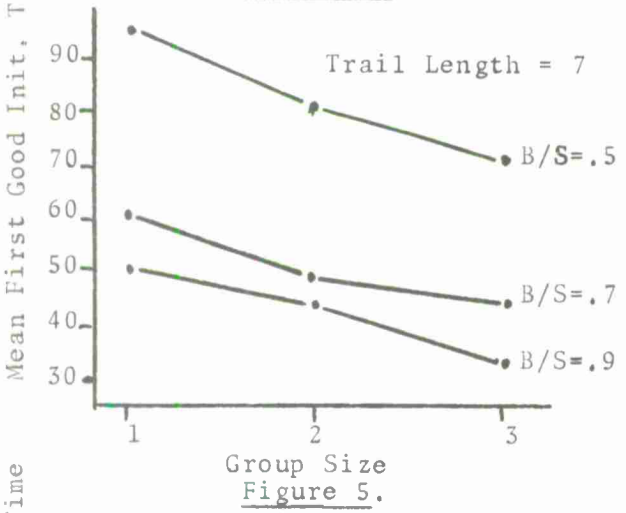



Figure 5.

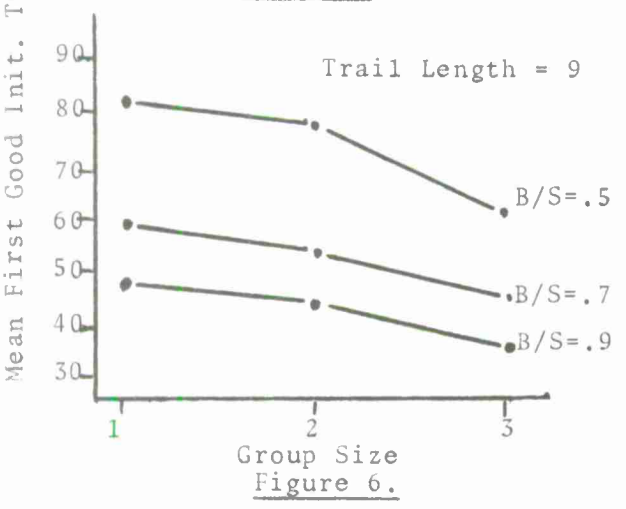



Figure 6.

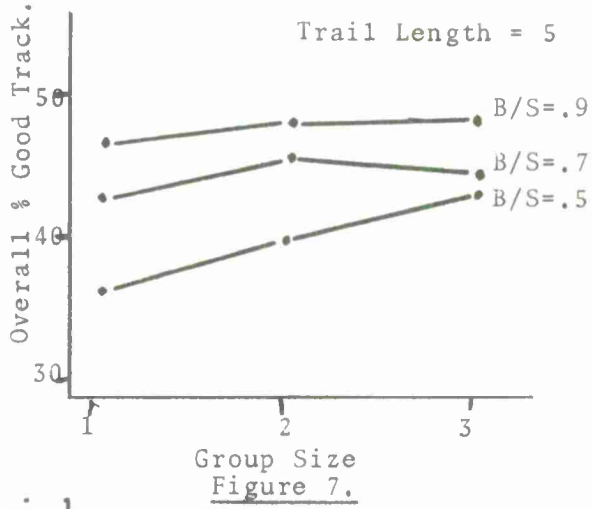



Figure 7.

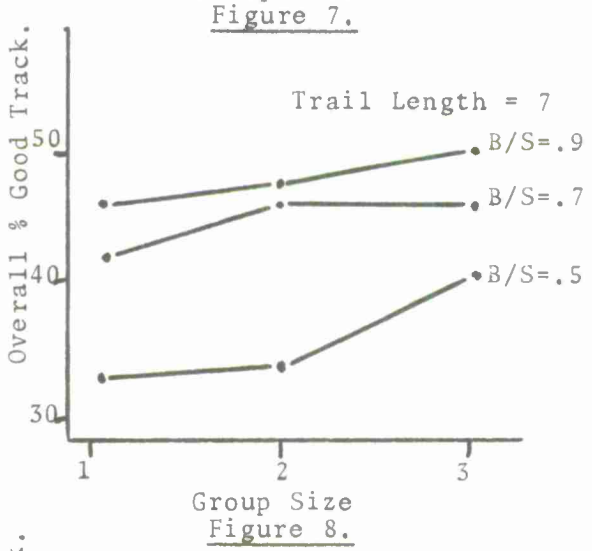



Figure 8.

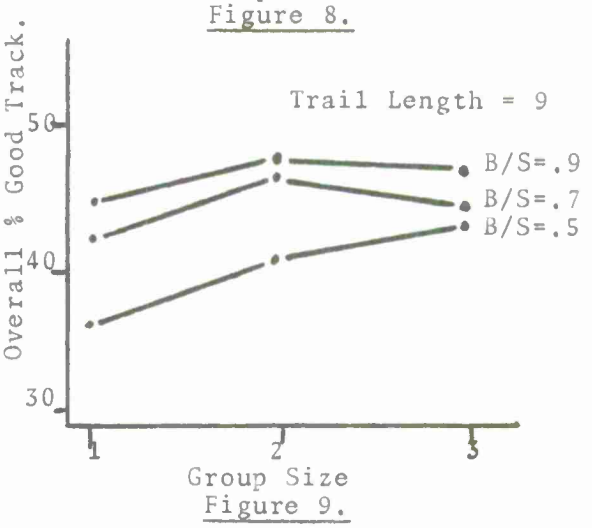



Figure 9.

Earlier in this report it was mentioned that shifts in operator strategy may be reflected in seemingly aberrant variability in the performance measures. Changing the composition of the team by varying the number of operators should also provoke a change in strategy. The quality of the information environment should also affect the way in which the task is approached. In Figure 10 the resulting shifts in trade-off between Percent Reinitiation Actions That Were Correct versus Percent Aircraft Correctly Initiated are portrayed. Note that when the information environment is poor (i.e., blip/scan = .5) the single operator seems to simply do the best he can and performance on these two measures is, accordingly, positively correlated. The tendency to trade off these two activities increases as the information environment improves, as evidenced by the negative correlations at the higher blip/scan values.

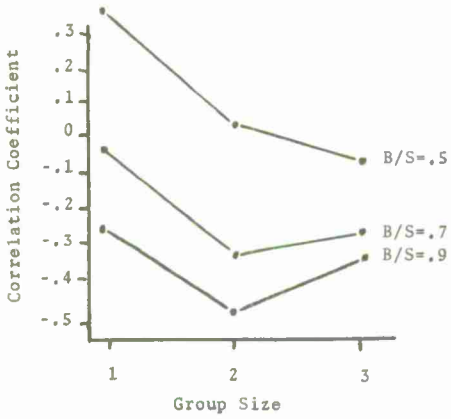



Figure 10. Percent Reinitiation Actions Correct versus Percent Aircraft Correctly Initiated.

Recommendations and conclusions were:

- Based on these preliminary data, two major engineering design recommendations can be made for digital radar surveillance systems. Both operator and surveillance system performance are predominantly driven by blip/scan ratio and not appreciably affected by trail length. Engineering efforts to increase blip/scan ratio should have significant payoffs in enhancement of surveillance system performance and ultimately overall operational effectiveness.

- The second recommendation concerns the assignment of surveillance areas to multiple surveillance technicians. On a cursory analysis, one might be led to conclude that a 2-man team might best split their functions such that one would be responsible for initiate actions and the other for reinitiate actions. This was the procedure adopted by the subjects in this study. Three operators might similarly partition their functions commensurate with the radar data environment of the moment. The performance data did support the hypothesis that multiple operators do generate improved surveillance system performance, but not to the extent one might anticipate. One reason for this may be that operators who monitor and perform only part of the sensing task required for initiation and maintenance are less likely to be aware of the total past and present activity of individual radar trails and have not been forced to attend the total situation. See Figure 11, comparing 1-, 2- and 3-man groups using a human performance measure of mean first good initiate time with a systems effectiveness measure of overall good tracking for different levels of probability of track failure.

- An alternative to the functional allocation of duties is, of course, to assign each operator his own (smaller) surveillance area and demand that he have total responsibility for all activities in that area.

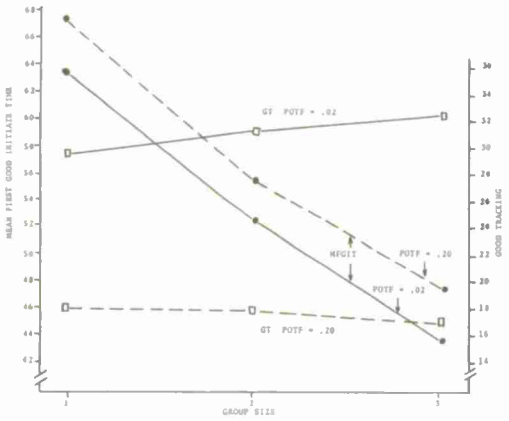



Figure 11. The Effect of Probability of Track Failure on MFGIT and GT

## THE OPERATOR AS A RESOURCE MANAGER

The ultimate goal, of course, of any command-control system is to commit deterrent or destructive resources which

most effectively minimize a developing threat. Assuming that the surveillance function achieves reasonable levels of target detection, acquisition, and tracking of the threat, the weapons direction function must be performed equally as effectively to win the encounter. The ability of the operator to interface with, gain access to, and manipulate information contained in computer storage is as important, or even more important, in the weapons direction/resource management function as it is in the surveillance function.

In one of the Air Force's new command-control systems, a critical design issue revolved around the most efficient means by which weapon directors could assign interceptors to an incoming threat of attacking bombers via interactive computer techniques.

Billy Crawford and William Pearson, from our group, recently developed a series of simulation efforts to evaluate the most effective interactive technique (computer control) for directing weapons under varying conditions of simulated threats.

Response Rate Study

Again the operator was stationed at our IBM 2250 consoles which were driven by the IBM 360-40H system. The console was modified to include a track ball control which controlled a cursor on the CRT. A diagrammatic representation of the console displays and controls is depicted in Diagram B.

There were three sets of experimental variables used in this study: the mode of control, the target introduction rate, and travel distance between a home point for a cursor and the interceptor/target locations. The mode of control included two versions of a light pen (LP-1 and LP-2) where LP-1 was the light pen normally found on the 2250 and LP-2, a modified version which achieved an optically sensitive area of approximately one mm. A third control mode was the track ball which controlled cursor movement on the scope. Target presentation rate was .75 seconds, 1.68 seconds, 2.85 seconds, and 4.0 seconds between successive targets introduced on the scope. The mean travel distance between cursor and target varied from .5 inches, 2 inches and 5 inches.

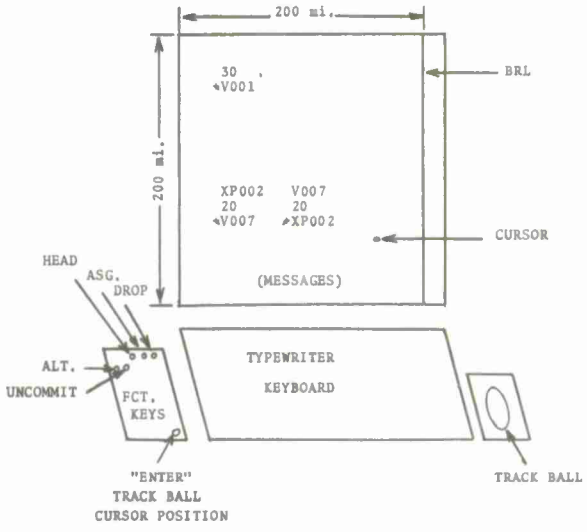



Diagram B. CRT Display and Controls for Weapons Direction Simulation

Dependent measures used in the study included probability of committing a response (subjects were instructed to only commit information assignments to the computer if he believed his cursor or light pen was aligned on the target), response time, and the distance between the target on which the subject's response registered.

The results are respectively represented for each of the measures in Tables V, VI, and VII. From Table V it can be seen that the operator is willing to commit a response with roughly a 30 percent greater probability for light pens than for track ball. Table VI shows that the track ball takes about 55-65 percent longer to commit responses and from Table VII the average final position error is more than twice as large with the track ball than the modified pen (LP-2).

These findings certainly verify one's intuitive human engineering judgment that a free or random access interactive control (light pen) provides superior data query and assignment capability than does a position constrained query device (track ball).

To verify these preliminary findings, a higher fidelity weapons direction simulation was developed.

TABLE V

AVERAGE PROBABILITY OF RESPONSE COMMITMENT FOR THE EXPERIMENTAL CONDITIONS

| Mode | Time Between Targets | Distance 0.5 | 2.0 | 5.0 | Average |
|---|---|---|---|---|---|
| LP-1 | .75 | .487 | .35 | .162 | .79 |
| | 1.68 | .912 | .912 | .862 | |
| | 2.85 | .9500 | .950 | .938 | |
| | 4.00 | .9750 | .962 | .975 | |
| LP-2 | .75 | .312 | .387 | .150 | .77 |
| | 1.68 | .900 | .925 | .875 | |
| | 2.85 | .950 | .962 | .925 | |
| | 4.00 | .962 | .950 | .912 | |
| TB | .75 | .262 | .212 | .212 | .50 |
| | 1.68 | .662 | .175 | .075 | |
| | 2.85 | .975 | .637 | .287 | |
| | 4.00 | 1.000 | .925 | .562 | |

TABLE VI

AVERAGE RESPONSE TIME IN SECONDS FOR THE EXPERIMENTAL CONDITIONS

| Mode | Time Between Targets | Distance 0.5 | 2.0 | 5.0 | Average |
|---|---|---|---|---|---|
| LP-1 | .75 | .691 | .684 | .739 | .974 |
| | 1.68 | .895 | .981 | 1.110 | |
| | 2.85 | .999 | 1.074 | 1.234 | |
| | 4.0 | 1.005 | 1.092 | 1.244 | |
| LP-2 | .75 | .673 | .684 | .717 | 1.142 |
| | 1.68 | .921 | 1.041 | 1.197 | |
| | 2.85 | 1.184 | 1.243 | 1.509 | |
| | 4.0 | 1.313 | 1.451 | 1.774 | |
| TB | .75 | .636 | .654 | .640 | 1.779 |
| | 1.68 | 1.291 | 1.519 | 1.571 | |
| | 2.85 | 1.743 | 2.341 | 2.676 | |
| | 4.0 | 1.894 | 2.863 | 3.524 | |

TABLE VII

AVERAGE DISTANCE IN INCHES BETWEEN TARGET AND RESPONSE MODE INDICATOR

| Mode | Time Between Targets | Distance 0.5 | 2.0 | 5.0 | Average |
|---|---|---|---|---|---|
| LP-2 | .75 | .50 | 1.28 | 4.34 | .71 |
| | 1.68 | .12 | .22 | .69 | |
| | 2.85 | .09 | .14 | .43 | |
| | 4.0 | .07 | .16 | .50 | |
| TB | .75 | .58 | 2.01 | 4.94 | 1.84 |
| | 1.68 | .31 | 1.85 | 4.94 | |
| | 2.85 | .08 | .93 | 3.79 | |
| | 4.0 | .07 | .29 | 2.25 | |

## Weapons Direction Simulation

Although not a complete representation of the weapons director's job in current command-control systems (i.e., BUIC III), the key tasks of interceptor-to-target assignments, heading change, altitude change, and commit-uncommit (drop) were simulated. Since one of the major questions dealt with track ball capability to perform these functions under peak threat load conditions, three levels of task loading were employed: (1) two waves of 8 fakers, each with 2 interceptors capable of kill; (2) one wave of 16 fakers, with 4 interceptors capable of kill; and (3) one wave of 20 fakers against 4 fighters. Faker speed was 500 knots and interceptor speed, 620 knots. Certainly one can argue that the threat/interceptor loads were not necessarily on an increasing ordinal scale, or even a unidimensional scale; however, they did achieve differing and perhaps realistic operational workload conditions.

Figure 12 summarizes the findings of this study using the average distance at which the threats were killed before their bomb release line (BRL). Most operational command-control personnel agree to the distance from BRL as an acceptable systems effectiveness measure assuming certain invasion geometry/time relationships.

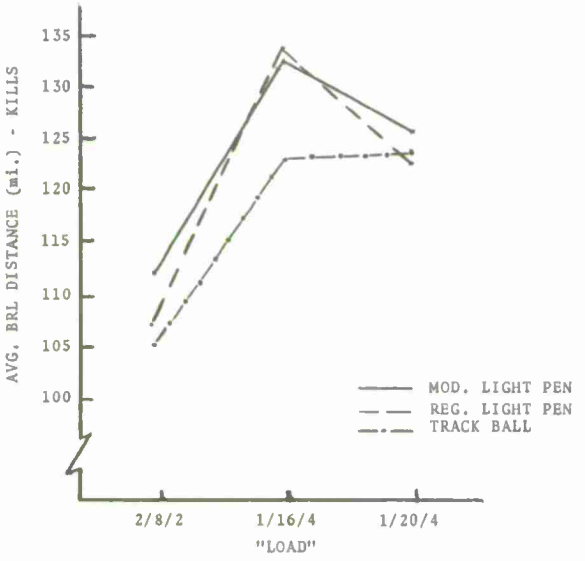



Figure 12. Summary

## SUMMARY

In this paper we have attempted to address problems of evaluating the effectiveness of several existing and advanced radar surveillance and command-control systems. This effectiveness depends largely on the ability of the operator to acquire a nonimaging and digitally displayed target and to identify what he perceives as being different from nontarget returns using the time course history of the returns. He must then assign this information to a computer programmed to automatically track the progress of the target across the surveilled area. The operator may perform this coordinate position identification and subsequent track initiation with a light pen or through cursor positioning with a track ball, joystick, or other means. The weapons direction function must be performed as effectively as the surveillance function in order to achieve maximum defense capability. The ability of the operator to efficiently interact with the computer-stored information about the acquired and tracked targets will depend, to a great extent, on the computer interactive device used.

Several simulation studies were reviewed which were aimed at assessing the operator's capabilities as a threat detector and resource manager under varying conditions of radar system performance and threat-environment dynamics.

Some generalizable principles have evolved from these simulations:

- The human operator is more effective than originally thought in acquiring targets in highly cluttered digital radar systems even under high information load conditions. That is, the $P_D$ (probability of penetrator detection) decreases from a maximum of 1.00 (all penetrators detected) for a target introduction rate of one/minute to a minimum of .9 for four penetrators/minute. A decrease of only 10 percent; the operator is an excellent perceptual filter.

Again from Figure 12, it appears that the track ball is inferior to the light pen versions in the order of 5-15 percent across "load" conditions--not as profound a difference as seen in the response time study. That is to say, the penetrators got closer to the bomb release line before destroyed under all load conditions with the track ball than with the light pens.

- The more effort the surveillance operator devotes to cleaning up uninitiated tracks on his display--keeping the computer updated for all existing tracks, the better the chance he has of picking up new incoming penetrators--confirms basic vigilance research findings.

- Simply adding surveillance operators may introduce coordination/communication requirements which increase as tracking system degradation increases and outweighs the seemingly positive effects of dividing the workload.

- Adding additional operators has little effect on overall good tracking of penetrators throughout a 60-minute invasion threat; however, to detect, initiate and eventually commit interceptor resources, more operators are a distinct advantage since they detect and initiate tracks sooner.

- As general human engineering principles might indicate, a free or random access computer interactive capability is superior for performing resource management functions to position-cursor dependent devices like track balls and joysticks.

## ACKNOWLEDGMENT

The research reported in this paper was conducted by personnel of the Aerospace Medical Research Laboratory, Aerospace Medical Division, Air Force Systems Command, Wright-Patterson Air Force Base, Ohio. This paper has been identified by Aerospace Medical Research Laboratory as AMRL-TR-72-104. Further reproduction is authorized to satisfy needs of the U.S. Government.

## REFERENCES

Howell, W. C. Some Principles for the Design of Decision Systems: A Review of Six Years of Research on a Command-Control System Simulation, AMRL-TR-67-136, Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio, September 1967.

Howell, W. C. and C. F. Gettys. Some Principles for Design of Decision Systems: A Review of the Final Phase of Research on Command-Control System Simulation, AMRL-TR-68-158, Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio, November 1968.

Mills, R. G. and M. A. Bauer. Aircraft Track Initiation and Maintenance on a Single-Operator Simulated Surveillance System: Technical Reports I and II, AMRL-TR-70-103 and AMRL-TR-71-76, respectively, Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio, July 1971 and April 1972, respectively.

Pearson, W. H. Time Simulation of an Air Surveillance Task With Varying Amounts of Radar Information, AMRL-TR-72-24, Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio, (in press).

Scanlon, L. A., S. N. Roscoe and R. C. Williges. "Time-Compressed Displays for Target Detection", Aviation Research Monographs, 1971, 1, 41-66.

## BIOGRAPHICAL SKETCHES

DR. H. H. BAILEY has degrees in Physics from Haverform College (AB) and the California Institute of Technology (PhD). After some teaching and a stint at the MIT Radiation Laboratory, Dr. Bailey joined the technical staff of the Bell Telephone Laboratories where he worked on radar bombing and navigation systems. In 1953 he went to Rand where he has worked on air-to-surface missiles for SAC and on a variety of guidance problems associated with TAC. Recently he has been actively studying navigation and weapon delivery techniques for tactical aircraft, night sensors for airborne use, and the performance of human observers.

LUCIEN M. BIBERMAN, a visiting professor of University of Rhode Island and a staff member in the Research and Engineering Support Division of IDA, began his career in 1942 as a physical chemist at NAIRN Research Laboratories. He was associated with the Naval Ordnance Test Station for 13 years as physicist in charge of the Physical Measurements Group; as head of the Ballistics Instrumentation Branch; as a consultant to associate technical director; and as a consultant to Aviation Ordnance Department and Weapons Development Department. In 1957 he became associate director of laboratories for Applied Sciences at the University of Chicago.

Mr. Biberman has authored and co-authored numerous books and journal articles. He is an active member of American Institute of Physics, American Physical Society, National Academy of Sciences, Optical Society of America, Institute of Radio Engineers, and the Institute of Electrical and Electronics Engineers.

DR. MARION R. BRYSON, scientific advisor, U. S. Combat Developments Experimentation Command at Ft. Ord, California received his PhD in statistics in 1958 from Iowa State College (now University), an MA in Mathematics in 1950 and a BS in Education in 1949 from University of Missouri.

From 1968-1972 he was Technical Director, Systems Analysis Group, U. S. Army Combat Developments Command, Ft. Belooir, Virginia. From 1958-1968 he was Director of Special Research in Statistics (Statistical and Operations Research Consultant to the U. S. Army Research Office - Durkam). In addition to these professional positions he has been an instructor of Mathematics at University of Missouri, University of Idaho, and Drake University of Des Moines, Iowa; at Duke University an associate professor of Mathematics, associate professor of Community Health Sciences, and Statistical Consultant in the School of Medicine.

Dr. Bryson has been affiliated with the following professional organizations: American Statistical Association (Faculty Representative, 1965-1968) (chapter secretary, 1967-1968), Institute of Mathematical Statistics, Operations Research Society of America (full member) (Program chairman, Military Applications Section, 1967-1968), Military Operations Research Society (Member, Board of Directors, 1970- 1st Vice President, 1972-).

RONALD A. ERICKSON is the head of the Human Factors Branch of the Naval Weapons Center, China Lake, California. He received a BS degree in physics from Idaho State College, studied one year on scholarship at the University of Munich, and obtained an MS degree in engineering from the University of California at Los Angeles in 1963. His work at the Center has included systems analysis on airborne weapon delivery systems, research and analysis on air-to-ground visual acquisition of targets, and target acquisition via television.

Mr. Erickson is the author of 5 and co-author of 8 U. S. Navy Technical Publications and has 6 open literature publications all dealing with Target Acquisition. He is a member of the Optical Society of America, the Human Factors Society, and the National Academy of Sciences Armed Forces Committee on Vision.

DR. MELVIN FREITAG has been Staff Engineer in the Human Factors Group at Martin Marietta Aerospace, Orlando, Florida for the past 2 1/2 years. He has been performing target acquisition studies both with man-in-the-loop and with completely computerized mathematical modeling approaches. For the past year he has specialized in the modeling and testing of a second generation FLIR in target acquisition air-to-ground missions as part of a team working under U. S. Army (Frankfort Arsenal) contract. At the present time Dr. Freitag is task leader on Air Force (AMRL) study to determine the effects of scene rotation on target acquisition and tracking performance.

Dr. Freitag holds a PhD in Psychology from the University of Virginia (1956), is a member of the Human Factors Society, and of the SPIE and has published over 30 reports and journal articles in the field of human factors, reliability and maintainability.

JAMES D. GILMOUR is manager of Attack Subsystems Technology Staff of the Boeing Company in Seattle where he is responsible for integration of the research, analysis, and systems engineering technologies required to insure an improved weapon system capability to identify and destroy enemy surface targets under all required operating conditions. He has worked on the prediction and improvement of attack system performance, and has conducted an extended series of studies on the acquisition, designation and tracking of ground targets by high performance aircraft under day, night and adverse weather conditions. As an Air Force pilot from 1954-59 he gained 5 years of tactical jet fighter experience in all-weather operations, radar fire control, maintenance flight test, and instrument crew training.

Mr. Gilmour received his BS degree from University of Wyoming in 1954, and his MS degree from Purdue University in 1959.

CARL P. GRAF, Research Engineer, maintains two vision laboratories, the first one containing a projection system which offers uniform intensity and precise control of target-to-background contrast; the second designed to evaluate the Honeywell Helmet Mounted Sight/Display system. In addition, he maintains the digital image processing system used in the reported study. Mr. Graf has developed a wide range of mechanical, optical, and electronic apparatus for conducting vision experiments. Mr. Graf is a contributing author to more than a dozen documents concerning optical systems and vision, is a Registered Professional Engineer (State of Minnesota), holds one design patent, and had one NASA invention disclosure.

DR. DANIEL B. JONES has had over 26 years experience in human factors engineering, operations research systems analysis, and military operations. Currently he is responsible for Human Factors for Martin Marietta Aerospace Orlando Division programs.

Prior to joining Martin Marietta in 1969, Dr. Jones was Chief of Human Factors Engineering for McDonnell-Douglas Astronautics Company Western Division. He spent 20 years as an Army officer engaged in human factors technology research and development and operations analysis.

Dr. Jones holds a PhD in psychology and statistics from the University of Missouri, Kansas City, an MS degree in experimental psychology from Tulane University, and a BAE degree (with honors) from the University of Florida.

ROBIN L. KEESEE received a BS in Inudstrail Engineering from Virginia Polytechnic Institute and State University in 1970. He is presently working toward completion of the PhD in Industrial Engineering and Operations Research at V. P. I., specializing in Human Factors. He is an associate member of the Human Factors Society.

DR. MARJORIE J. KREBS is a principal research scientist at Honeywell, Inc. where she is involved in reserach related to vision and visual displays. Before coming to Honeywell she was an assistant professor of psychology at New Mexico State University where this study was conducted. Prior to completing her doctoral work at the University of Virginia she worked at the American Institutes for Research and at HRB-Singer, Inc.

SHELDON H. LEVINE joined McDonnel Douglas in 1966, where he has worked primarily in the areas of reconnaissance and electronic display systems with emphasis on visual process and target acquisition. Over the past 3 years he was responsible for the human factors efforts in the imagery exploitation program of the McDonnell Douglas Reconnaissance. He is currently working on task and decision analyses relating to management functions of the G-2 Air Officer. Mr. Levine received an

MA in experimental psychology in 1964 from Adelphi University, Garden City, New York.

WAYNE L. MARTIN is an engineering research psychologist in the Human Engineering Division of the Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio. He received his BA degree (1964) and MS degree (1965) in psychology at the Eastern Washington State College. Post graduate work includes two years in experimental psychology at the University of Kentucky and a one-year program in which he is presently enrolled at Purdue University. During the past five years, he has conducted research in the areas of maintainability, logistics management, visually coupled systems and command, control and surveillance systems.

RALPH L. MCCLUSKEY is currently the Chief of Development for the Army's Electromagnetic Environmental Test Facility (EMETF) - a facility of the US Army Electronic Proving Ground at Fort Huachuca, Arizona. He was formerly with the US Air Force Development Laboratories and Strategic Air Command. He left the Air Force to become Technical Director for the Communications Electronics Agency of the US Army Combat Development Command. When the Communications Electronics Agency moved from Fort Huachuca to Fort Monmouth, New Jersey, Mr. McCluskey transferred to EMETF as Chief of Development.

DR. ANTONIO J. MENDEZ is a Senior Staff Engineer in the Electro-Optics Department of Martin Marietta Aerospace Orlando, Florida. For the past four years he has been task leader for infrared research, primarily concerned with the optimization of IR sensor systems for day/night target acquisition and fire control. He supervises the Infrared Technology Section consisting of an interdisciplinary group of scientists, engineers, mathematicians and technicians. In this position he is responsible for infrared systems analysis on proposed fire control systems, mathematical modeling of FLIR and other target acquisition sensors, research and development of advanced equipment and their testing by simulation and field tests.

Dr. Mendez holds a PhD degree in Physics (1968) from USC and is a member of the Phi Beta Kappa and the American Physical Society.

DR. JOHN J. O'HARE, with special interest in performance effectiveness programs that enhance perceptual and cognitive functions as well as extend the theoretical bases and refine behavior models, has been employed as a research psychologist within the Office of Naval Research Engineering Psychology Programs since June 1971. These interests have been developed and extended in prior employments which included the U.S. Navy Medical Research Lab, International Electric Corp., United Aircraft Corp., MITRE Corp., and Georgetown University.

FRED A. PAYNE, Vice President and Chief Engineer, Martin Marietta Aerospace, Orlando is responsible for all Orlando Division Engineering activities, Logistics activities, Facilities Planning, and the Research and Technology Program.

In February of 1968, Mr. Payne came to Orlando as Director of Technical Operations from the Marquardt Corporation where he was Vice President - Corporate Development. Before joining Marquardt in 1965, he served four years in the Department of Defense as Deputy Director of Defense Research and Engineering, and as Assistant Director of Defense Research and Engineering for Strategic Weapons. For the previous 20 years he had been associated with North American Aviation. In 1966, Mr. Payne took a leave of absence from Marquardt, to serve as Director of the STRAT-X Study Program at the Institute for Defense Analyses.

A native of Washington, D. C., Mr. Payne holds a degree in chemistry from the University of Washington, Seattle.

DR. JAMES J. REGAN is Chief Psychologist, Head of the Human Factors Laboratory and Director of the Training Analysis and Evaluation Group at the Naval Training Equipment Center. His graduate degrees were obtained at Fordham University. He is a Fellow of the American Psychological Association and Senior Member of the Institute for Electrical and Electronic Engineers.

FREDERICK A. ROSELL is an Advisory Engineer with the Westinghouse Electric Corporation Defense and Electronic Systems Center in Baltimore, Maryland. A graduate of Rensselaer Polytechnic Institute in 1953, Mr. Rosell has been active in the design and synthesis of infrared and low light television systems since that time. His most recent efforts have been concentrated in the design of imaging systems through analytical modeling of sensors and parallel psychophysical experimentation to determine

observer requirements for target detection and recognition.

Mr. Rosell is a member of the Optical Society of America, the IEEE working group on silicon camera tubes, and the IRIS working group on imaging devices.

DR. ALVIN D. SCHNITZLER received his PhD in Solid State Physics from the University of Maryland in 1961. He was employed at the Naval Research Laboratory as a physicist from 1952 to 1962, at Melpar, Inc. as a senior scientist from 1962 to 1963, and at the U. S. Army Night Vision Laboratory as a supervisory physicist from 1963 to 1969. Since 1969 Dr. Schnitzler has been a member of the staff of the Institute for Defense Analyses where he is engaged in studies related to the development and military applications of electro-optical systems.

Dr. Schnitzler serves as a member of the Working Group on Special Devices of the Advisory Group on Electron Devices to the Office of the Director of Defense Research and Engineering, and is a member of APS, OSA and IEEE.

DR. HERSCHEL C. SELF is a Research Psychologist with the 657th Aerospace Medical Research Laboratory at Wright-Patterson Air Force Base, Ohio. For over a decade he and his co-workers at the lab have been examining target detection and recognition via laboratory simulations and from aircraft. The unaided eye as well as a large variety of sensors and displays have been used. Prior to this period he was an engineering psychologist at the Naval Research Laboratory.

Interests and hobbies of Dr. Self include experimental psychology, vision, optics, amateur astronomy, electronics, and table tennis.

DR. HARRY L. SNYDER received his AB degree from Brown University, and his MA and PhD degrees from The Johns Hopkins University in experimental psychology. He has held research and management positions at Autonetics/North American Rockwell and the Boeing Company, where he dealt primarily with problems of visual displays, sensors, and image quality. As Professor of Industrial Engineering and Operations Research, and Director of the Human Factors Laboratory at V. P. I., his current research interests are in the areas of television system design and image quality, pseudocolor, and transportation system command/control.

WARREN H. TEICHNER is professor of psychology at New Mexico State University, Las Cruces. Previous positions were at the American Institutes for Research, Northeastern University, Tufts University, the University of Massachusetts, the U. S. Army Natic Laboratories, and the Aeromedical Laboratory, WADC. For the past 14 years he was also consultant investigator at the Guggenheim Center for Aerospace Health and Safety, Harvard University School of Public Health.

DR. MARTIN A. TOLCOTT, Director of the Engineering Psychology Programs, Psychological Sciences Division, Office of Naval Research, is responsible for a program of research in man-machine problems which includes human engineering and performance effectiveness. He is a Fellow of the American Psychological Association and the Human Factors Society; he served as President and is currently a Director-at-Large of the Potomac Chapter, Human Factors Society.

DR. DONALD A. TOPMILLER is the chief of the Systems Effectiveness Branch, Human Engineering Division, Aerospace Medical Research Laboratory at Wright-Patterson Air Force Base, Ohio. He received a BA degree in general psychology from Miami University, an MA degree in engineering psychology from Lehigh University, and a PhD degree in industrial psychology and statistics from Ohio State University in 1964. He has been employed by the Aerospace Medical Research Laboratory since 1956 and has served as a lecturer in psychology at Wright State University since 1965. His research interests are in the areas of man-computer systems simulation and human factors, and systems effectiveness, including maintainability and human performance reliability quantification.

DR. LEON G. WILLIAMS, Senior Principal Research Scientist, Honeywell Systems and Research Division, Minneapolis, Minnesota, has been studying the acquisition of information from displays since joining Honeywell in 1961. His main research interests are in the use of computers to process and display information and in the visual response to displayed information. He has developed mathematical models of the search process based on several studies of eye movements during visual search. These have led to a procedure for determining optimum coding

of information in displays. He has designed optical and electronic equipment for measuring and recording eye movements and for controlling related apparatus. He has also designed and validated an automatic character recognition technique. This latter led to the discovery of a novel method of secure communication. Prior to joining Honeywell, Dr. Williams taught courses in psychology at Eastern Michigan University and conducted experiments relating to display design at Dunlap and Associates.

DR. ROBERT H. WILLSON received the PhD degree from the University of Illinois, Urbana campus in 1966. Since then, he has been employed by Westinghouse Electric Corporation at the Baltimore Defense and Electronic Systems Center. He is currently a Fellow Engineer in the Electro-Optical Systems Development section where he does analysis and psychophysical experiments appropriate for IR and LLLTV systems.

DR. EDWARD W. YOUNGLING held a variety of positions within the aircraft and astronautics companies in St. Louis since joining McDonnell Douglas in 1965. At the present time he is a Section Manager in the Engineering Psychology Department. Prior to joining McDonnell, he was a Research Psychologist at the U. S. Army Natick Laboratories in Massachusetts. From 1959 until 1962 he was a Research Assistant to Dr. Warren H. Teichner at his Institute of Environmental Psychophysiology at the University of Massachusetts in Amhearst.

DISTRIBUTION LIST FOR UNCLASSIFIED VOLUME I OF
TECHNICAL REPORT ON PROCEEDINGS OF TARGET
ACQUISITION SYMPOSIUM, 14, 15, 16 NOVEMBER 1972.

Director, Engineering Psychology
Programs, Code 455
Office of Naval Research
800 North Quincy Street
Arlington, Virginia 22217

Defense Documentation Center
Cameron Station
Alexandria, Virginia 22314

Director, ONR Branch Office
Attn: Dr. C. Harsh
495 Summer Street
Boston, Massachusetts 02210

Director, ONR Branch Office
Attn: Dr. M. Bertin
536 S. Clark Street
Chicago, Illinois 60605

Director, ONR Branch Office
Attn: Dr. E. Gloye
1030 East Green Street
Pasadena, California 91106

Director, ONR Branch Office
Attn: Mr. R. Lawson
1030 East Green Street
Pasadena, California 91106

Director, Naval Research Laboratory
Technical Information Division
Code 2027
Washington, D. C. 20375

Director, Naval Research Laboratory
Attn: Library, Code 2029 (ONRL)
Washington, D. C. 20375

Mr. John Hill
Naval Research Laboratory
Code 5634
Washington, D. C. 20375

Office of Naval Research
Aeronautics Programs, Code, 461
Department of the Navy
Arlington, Virginia 22217

Office of Naval Research
Code 461
Attn: JANAIR Chairman
Arlington, Virginia 22217

Dr. John J. Collins
Office of the Chief of Naval
Operations, Op-987F
Department of the Navy
Washington, D. C. 20350

CDR H. J. Connery
Office of the Chief of Naval
Operations, Op-987M4
Department of the Navy
Washington, D. C. 20350

Dr. A. L. Slafkosky
Scientific Advisor
Commandant of the Marine Corps
Code AX
Washington, D. C. 20380

Dr. Heber G. Moore
Hqs., Naval Material Command
Code 03R4
Department of the Navy
Washington, D. C. 20360

Commander, Naval Air Systems Command
Crew Systems Division, AIR 531
Washington, D. C. 20360

Commander, Naval Air Systems Command
Attn: Mr. J. Wolin, AIR-53371
Washington, D. C. 20360

Commander, Naval Air Systems Command
Attn: Mr. T. Momiyama, AIR 3034
Washington, D. C. 20361

Commander, Naval Electronics
Systems Command
Command and Display Systems Branch
Code 0544
Washington, D. C. 20360

Mr. Joseph B. Blankenheim
Naval Electronics Systems Command
Code 0474
Washington, D. C. 20360

Commander, Naval Air Systems Command
NAVAIR 340F
Washington, D. C. 20361

Mr. James Jenkins
Naval Ships Systems Command
Code PMS 302-43
Washington, D. C. 20360

CDR James E. Goodson
Bureau of Medicine and Surgery
Operational Psychology Branch, Code 513
Department of the Navy
Washington, D. C. 20372

Mr. A. Sjoholm
Bureau of Personnel
Personnel Research Div., PERS A-3
Washington, D. C. 20370

Commander Naval Safety Center
Naval Air Station
Norfolk, Virginia 32511

LCDR Curt Sandler, MSC
Naval Safety Center
Code 811
Norfolk, Virginia 23511

CDR Robert Wherry
Human Factors Engineering Systems Ofc.
Naval Air Development Center
Johnsville
Warminster, Pennsylvania 18974

Human Factors Engineering Branch
Code 5342
U. S. Naval Missile Center
Point Mugu, California 93041

Aeromedical Branch
Service Test Division
U. S. Naval Air Test Center
Patuxent River, Maryland 20670

Mr. Ronald A. Erickson
Head, Human Factors Branch, Code 4011
Naval Weapons Center
China Lake, California 93555

Human Engineering Branch, Code A624
Naval Ship Research and Development Center
Annapolis Division
Annapolis, Maryland 21402

Dr. Robert French
Naval Undersea Center
San Diego, California 92132

Mr. Richard Coburn
Head, Human Factors Disision
Naval Electronics Laboratory Center
San Diego, California 92152

Dean of Research Administration
Naval Postgraduate School
Monterey, California 93940

Mr. E. Ramras
Technical Director
Personnel Research and Development Lab
Washington Navy Yard
Washington, D. C. 20374

Commanding Officer
Naval Personnel and Training
Research Laboratory
Attn: Technical Director
San Diego, California 92152

CAPT Allen McMichael
Chief of Naval Training (Code 017)
Naval Air Station
Pensacola, Florida 32508

Dr. J. J. Regan
Human Factors Department, Code N215
Naval Training Equipment Center
Orlando, Florida 32813

Dr. George Moeller
Head, Human Factors Engineering Branch
Submarine Medical Research Laboratory
Naval Submarine Base
Groton, Connecticut 06340

CDR Thomas Gallagher
Chief, Aerospace Psychology Division
Naval Aerospace Medical Institute
Pensacola, Florida 32512

Dr. James T. Lester
Office of Naval Research Branch Office,
London
Box 39
FPO New York 09510

U. S. Air Force Office of
Scientific Research
Life Sciences Directorate, NL
1400 Wilson Blvd.
Arlington, Virginia 22209

Dr. J. M. Christensen
Chief, Human Engineering Division
Aerospace Medical Research Lab
Wright-Patterson AFB, Ohio 45433

Dr. Walter F. Grether
Behavioral Science Laboratory
Aerospace Medical Research Lab
Wright-Patterson AFB, Ohio 45433

Dr. J. E. Uhlaner
Director, U. S. Army Research Institute
for the Social and Behavioral Sciences
1300 Wilson Blvd.
Arlington, Virginia 22209

Chief of Research and Development
Human Factors Branch
Behavioral Science Division
Department of the Army
Washington, D. C. 20310
Attn: Mr. J. Barber

Technical Director
U. S. Army Human Engineering Labs
Aberdeen Proving Ground
Aberdeen, Maryland 21005

Lt Col Austin W. Kibler
Director, Behavioral Sciences
Advanced Research Projects Agency
1400 Wilson Blvd.
Arlington, Virginia 22209

U. S. Army, Combat Developments
Experimentation Command
Plans and Programs Office
Fort Ord, California 93941

Dr. Stanley Deutsch
Chief, Man-Systems Integration
OART, Hqs., NASA
600 Independence Avenue
Washington, D. C. 20546

Dr. Jesse Orlansky
Institute for Defense Analyses
400 Army-Navy Drive
Arlington, Virginia 22202

Dr. W. H. Teichner
Department of Psychology
New Mexico State University
Las Cruces, New Mexico 88001

Dr. L. J. Fogel
Decision Science, Inc.
4508 Mission Bay Drive
San Diego, California 92112